Discrete time approximation of decoupled Forward-Backward SDE driven by pure jump Lévy-processes

Soufiane Aazizi *

October 25, 2011

Abstract

We present a new algorithms to discretize a decoupled forward backward stochastic differential equations driven by pure jump Lévy process (FBSDEL in short). The method is built in two steps. Firstly, we approximate the FBSDEL by a forward backward stochastic differential equations driven by a Brownian motion and Poisson process (FBSDEBP in short), in which we replace the small jumps by a Brownian motion. Then, we prove the convergence of the approximation when the size of small jumps ε goes to 0. In the second step, we obtain the L^p Hölder continuity of the solution of FBSDEBP and we construct two numerical schemes for this FBSDEBP. Based on the L^p Hölder estimate, we prove the convergence of the scheme when the number of time steps n goes to infinity. Combining these two steps leads to prove the convergence of numerical schemes to the solution of FBSDEL.

Key words : Discrete-time approximation, Euler scheme, decoupled forward-backward SDE with jumps, Small jumps, Malliavin calculus.

MSC Classification: 60H35, 60H07, 60J75

*Department of Mathematics, Faculty of Sciences Semlalia Cadi Ayyad University, B.P. 2390 Marrakesh, Morocco. aazizi.soufiane@gmail.com

1 Introduction and summary

In this paper, we are concerned by discretization of a system of decoupled forward-backward stochastic differential equation (FBSDEs in short) driven by a pure jump Lévy process

$$\begin{cases} X_t = X_0 + \int_0^t b(X_r)dr + \int_0^t \int_{\mathbb{R}} \beta(X_{r-})\bar{M}(de,dr), \\ Y_t = g(X_T) + \int_t^T f(\Theta_r)dr - \int_t^T \int_{\mathbb{R}} V_r \bar{M}(de,dr) \end{cases} \tag{1.1}$$

Here $\Theta := \big(X, Y, \int_E \rho(e)Ve\nu(de)\big)$ and $\bar{M}(E,t) = \int_{E\times[0,t]} e\bar{\mu}(de,dr)$ where $\bar{\mu}(de,dr) := \mu(de,dr) - \nu(de)dr$ an independent compensated Poisson measure and μ a Poisson random measure on $\mathbb{R}\times[0,T]$ with intensity ν satisfying $\int 1\wedge|e|^2\nu(de) < \infty$.

Numerical discretization schemes for FBSDE have been studied by many authors. In the no-jump case, Ma et al. [21] developed the first step algorithm to solve a class of general forward-backward SDE. Douglas et al. [14] suggest a finite difference approximation of the associated PDE. Other discrete scheme have been considered in [7], [8] and [11] mainly based on approximation of the Brownian motion by some discrete process, Gobet et al. [18] proposed an adapted Longstaff and Schwartz algorithm based on non-parametric regressions. In the jump case, to our knowledge, there is only the work of Bouchard and Elie [5] in which the authors propose a Monte-Carlo methods in the case when $\nu(\mathbb{R}) < \infty$.

The main motivation to study the numerical scheme of a systems of above form, is to treat the case when $\nu(\mathbb{R}) = \infty$, which means the existence of an infinite number of jumps in every interval of non-zero length a.s.. In this sense, we should mention the important work on the approximation of stochastic differential equation studied by Kohatsu-Higa and Tankov [20].

Since we are interested in the case of $\nu(\mathbb{R}) = \infty$, we will follow the idea of [20] to approximate (1.2) without cutoff the small jumps smaller than ε, which should improve the approximation scheme. Then by using the approximation result of Asmussen and Rosinski [2] we replace the small jumps of the driven-Lévy process with $\sigma(\varepsilon)W$ where W is a standard Brownian motion and $\sigma^2(\varepsilon) := \int_{E^\varepsilon} e^2\nu(de)$.

In the aim to approximate (1.1), we cut the jumps at ε as the following

$$\begin{cases} X_t = X_0 + \int_0^t b(X_r)dr + \int_0^t \beta(X_{r-})dR_r + \int_0^t \int_{E_\varepsilon} \beta(X_{r-})\bar{M}(de,dr) \\ Y_t = g(X_T) + \int_t^T f(\Theta_r)dr - \int_t^T V_r dR_r - \int_t^T \int_{E_\varepsilon} V_r \bar{M}(de,dr) \end{cases} \tag{1.2}$$

where $R_t = \int_0^t \int_{|e|\leq\varepsilon} e\bar{M}(de,dr)$, $E^\varepsilon := \{e\in\mathbb{R},\ \text{s.t} \ /\ |e|\leq\varepsilon\}$, $E_\varepsilon := \{e\in\mathbb{R},\ \text{s.t} \ /\ |e|>\varepsilon\}$ and $E := \mathbb{R} = E^\varepsilon \cup E_\varepsilon$.

The idea we propose is to discretize the solution of (1.1) in two steps. In the first step, we approximate (1.2) by the following FBSDE:

$$\begin{cases} X_t^\varepsilon = X_0^\varepsilon + \int_0^t b(X_r^\varepsilon)dr + \int_0^t \beta(X_r^\varepsilon)\sigma(\varepsilon)dW_r + \int_0^t \int_{E_\varepsilon} \beta(X_{r-}^\varepsilon)\bar{M}(de,dr) \\ Y_t^\varepsilon = g(X_T^\varepsilon) + \int_t^T f(\Theta_r^\varepsilon)dr - \int_t^T Z_r^\varepsilon dW_s - \int_t^T \int_{E_\varepsilon} U_r^\varepsilon(e)\bar{M}(de,dr) \end{cases} \tag{1.3}$$

Here $\Theta^\varepsilon := \left(X^\varepsilon, Y^\varepsilon, \Gamma^\varepsilon\right)$ and $\Gamma^\varepsilon := \int_{E_\varepsilon} \rho(e) U^\varepsilon(e) e\nu(de)$. Further, we show that for a finite measure m defined by $m(E) := \int_E e^2 \nu(de)$, our error

$$\begin{aligned} Err^2_\varepsilon(Y,V) \;&:=\; \mathbb{E}\left[\sup_{t\le T}|Y_t - Y^\varepsilon_t|^2\right] + \mathbb{E}\left[\sup_{t\le T}\left|\int_0^t V_r dR_r - \int_0^t Z^\varepsilon_r dW_r\right|^2\right] \\ &+\mathbb{E}\left[\int_0^T \int_{E_\varepsilon} |V_r - U^\varepsilon_r(e)|^2 m(de) dr\right], \end{aligned}$$

is controlled by $\sigma(\varepsilon)^2$, which means that the solution of (1.3) converges to the solution of (1.1), as the size of small jumps ε goes to 0 (See Remark 2.1). We also derive the upper bound

$$\mathbb{E}\left[\sup_{t\le T}|X_t - X^\varepsilon_t|^2\right] \le C\sigma(\varepsilon)^2. \tag{1.4}$$

The second step consists of discretizating the approximated FBSDE (1.3) and studying its convergence to (1.2). For this purpose we consider two numerical schemes, the first one is based on discrete-time approximation of decoupled FBSDE derived by Bouchard and Elie [5]. More precisely, for a fixed ε, given a regular grid $\pi = \{t_i = iT/n, i = 0,1,...,n.\}$, the authors approximate X^ε by its Euler scheme $\bar{X}^\pi$ and $(Y^\varepsilon, Z^\varepsilon, \Gamma^\varepsilon)$ by the discrete-time process $(\bar{Y}^\pi_t, \bar{Z}^\pi_t, \bar{\Gamma}^\pi_t)$

$$\left\{\begin{array}{lcl} \bar{X}^\pi_{t_i+1} &=& \bar{X}^\pi_{t_i} + \frac{1}{n} b(\bar{X}^\pi_{t_i}) + \beta(\bar{X}^\pi_{t_i})\sigma(\varepsilon)\Delta W_{i+1} + \int_{E_\varepsilon} \beta(\bar{X}^\pi_{t_i}) \bar{M}(de,(t_i,t_{i+1}]) \\ \bar{Z}^\pi_t &=& n\mathbb{E}\left[\bar{Y}^\pi_{t_{i+1}} \Delta W_{i+1}/\mathcal{F}_{t_i}\right] \\ \bar{\Gamma}^\pi_t &=& n\mathbb{E}\left[\bar{Y}^\pi_{t_{i+1}} \int_{E_\varepsilon} \rho(e)\bar{M}(de,(t_i,t_{i+1}])/\mathcal{F}_{t_i}\right] \\ \bar{Y}^\pi_t &=& \mathbb{E}\left[\bar{Y}^\pi_{t_{i+1}}/\mathcal{F}_{t_i}\right] + \frac{1}{n} f(\bar{X}^\pi_{t_i}, \bar{Y}^\pi_{t_i}, \bar{\Gamma}^\pi_{t_i}) \end{array}\right. \tag{1.5}$$

on each interval $[t_i, t_{i+1})$, where the terminal value $\bar{Y}^\pi_{t_n} := g(\bar{X}^\pi_{t_n})$. Under Lipschitz continuity of the solution, the authors proved that the discretization error

$$\overline{Err}^2_n(Y^\varepsilon, Z^\varepsilon, \Gamma^\varepsilon) \;:=\; \sup_{t\le T}\mathbb{E}\left[|Y^\varepsilon_t - \bar{Y}^\pi_t|^2\right] + \int_0^T \mathbb{E}\left[|Z^\varepsilon_t - \bar{Z}^\pi_t|^2 + |\Gamma^\varepsilon_t - \bar{\Gamma}^\pi_t|^2\right]dt \tag{1.6}$$

achieves the optimal convergence rate $n^{-1/2}$. Finally, we derive the first main result of this paper in Proposition 3.1 showing that the approximation-discretization error

$$\overline{Err}^2_{(n,\varepsilon)}(Y,V) \;:=\; \sup_{t\le T}\mathbb{E}\left[|Y_t - \bar{Y}^\pi_t|^2 + \left|\int_0^t V_r dR_r - \int_0^t \bar{Z}^\pi_r dW_r\right|^2\right] + \|\Gamma - \bar{\Gamma}^\pi\|^2_{H^2} \tag{1.7}$$

is bounded by $C(n^{-1} + \sigma(\varepsilon)^2)$ and converges to 0 as (n,ε) tends to $(\infty, 0)$, where $\Gamma := \int_{E_\varepsilon} \rho(e) V e\nu(de)$. Taking $\varepsilon = n^{-1/2}$, our approximation-discretization achieves the optimal convergence rate $n^{-1/2}$.

The second numerical scheme has been inspired from the paper of Hu, Nualart and Song [19]. Where the authors study a backward stochastic differential equation driven by a Brownian motion with general terminal variable ξ. They propose a new scheme using the

representation of Z^ε as the trace of the Malliavin derivatives of Y^ε. Their discretization scheme is based on the L^p-Hölder continuity of the solution Z^ε, to obtain an estimate of the form

$$\mathbb{E}|Z_t^\varepsilon - Z_s^\varepsilon|^p \leq K|t-s|^{\frac{p}{2}},$$

which implies the existence of a γ-Hölder continuous version of the process Z^ε for any $\gamma < \frac{1}{2} - \frac{1}{p}$. In this sense, our article extend the work done in [19] to a forward-backward stochastic differential equation with jumps and terminal value $g(X_T^\varepsilon)$. Similarly to [19], we obtain the following regularity of Γ^ε

$$\mathbb{E}|\Gamma_t^\varepsilon - \Gamma_s^\varepsilon|^p \quad \leq \quad C|t-s|^{\frac{p}{2}},$$

which allows us to deduce the existence of a γ-Hölder continuous version of the process Γ^ε for any $\gamma < \frac{1}{2} - \frac{1}{p}$. Finally, on one hand, we use the representation of Z^ε and Γ^ε as the trace of Malliavin derivative of Y^ε to derive our a new extended discretization scheme for the solution $(Y^\varepsilon, Z^\varepsilon, \Gamma^\varepsilon)$ of (1.3). From other hand we approximate X^ε by X^π the continuous-time version of the Euler scheme, that is for a fixed $\varepsilon > 0$

$$\begin{cases} X_t^\pi &= X_{\phi_t^n}^\pi + \sigma(\varepsilon) b(X_{\phi_t^n}^\pi)(t-\phi_t^n) + \sigma(\varepsilon)\beta(X_{\phi_t^n}^\pi)(W_t - W_{\phi_t^n}) + \int_{E_\varepsilon} \beta(X_{\phi_t^n}^\pi)\bar{M}(de,(t,\phi_t^n]). \\ Y_{t_i}^\pi &= \mathbb{E}\Big[Y_{t_{i+1}}^\pi + f(\Theta_{t_{i+1}}^\pi)\Delta t_{i+1}/\mathcal{F}_{t_i}\Big] \\ Z_{t_i}^\pi &= \mathbb{E}\Big[\mathcal{E}_{t_{i+1},t_n}^\pi \partial_x g(X_T^\pi) D_{t_i} X_T^\pi + \sum_{k=i}^{n-1} \mathcal{E}_{t_{i+1},t_{k+1}}^\pi \partial_x f(\Theta_{t_{k+1}}^\pi) D_{t_i} X_{t_k}^\pi \Delta t_k / \mathcal{F}_{t_i}\Big] \\ \Gamma_{t_i}^\pi &= \mathbb{E}\Big[\int_{E_\varepsilon} \rho(e)\big[\mathcal{E}_{t_{i+1},t_n}^{e,\pi} D_{t_i,e} g(X_T^\pi) + \sum_{k=i}^{n-1} \mathcal{E}_{t_{i+1},t_{k+1}}^{e,\pi} \alpha_{t_i,t_{k+1}}^\pi D_{t_i,e} X_{t_{k+1}}^\pi \Delta t_k\big]\nu(de)/\mathcal{F}_{t_i}\Big] \end{cases} \tag{1.8}$$

with terminal values $Y_{t_n}^\pi = g(X_T^\pi)$, $Z_{t_n}^\pi = \sigma(\varepsilon)\partial_x g(X_T^\pi)\beta(X_T^\pi)$ and $U_{t_n,e}^\pi = g(X_T^\pi + \beta(X_T^\pi)) - g(X_T^\pi)$, where ϕ_t^n, $\mathcal{E}_{t_i,t_j}^\pi$ and $\mathcal{E}_{t_i,t_j}^{e,\pi}$ are detailed in section 4.

The key-ingredient for computation of discretization error, is based on the L^p-Hölder continuity of the solution $(Y^\varepsilon, Z^\varepsilon, \Gamma^\varepsilon)$. This allows us to prove that

$$Err_n^2(Y^\varepsilon, Z^\varepsilon, \Gamma^\varepsilon) \quad := \quad \mathbb{E}\max_{0\leq i\leq n}\Big[|Y_{t_i}^\varepsilon - Y_{t_i}^\pi|^2 + |Z_{t_i}^\varepsilon - Z_{t_i}^\pi|^2 + |\Gamma_{t_i}^\varepsilon - \Gamma_{t_i}^\pi|^2\Big],$$

is controlled by $|\pi|^{1-\frac{1}{log\frac{1}{|\pi|}}}$. Then we obtain the second main result of this article in Theorem 4.3, which stating that

$$\begin{aligned} Err_{n,\varepsilon}^2(Y,V) \quad := \quad & \max_{0\leq i\leq n} \sup_{t\in[t_i,t_{i+1}]} \mathbb{E}\left[|Y_t - Y_{t_i}^\pi|^2\right] + \mathbb{E}\left|\int_0^T V_r dR_r - \sum_{i=0}^{n-1} Z_{t_i}^\pi \Delta W_{t_i}\right|^2 \\ & + \sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}} \mathbb{E}|\Gamma_t - \Gamma_{t_i}^\pi|^2 dt, \end{aligned}$$

is of the order $\sigma(\varepsilon)^2 + |\pi|^{1-\frac{1}{log\frac{1}{|\pi|}}}$ and converges to 0 as the discretization step (ε, n) tends to $(0,\infty)$.

The importance of the above scheme, is it can be adapted to the case when the a terminal value is not given by the forward diffusion equation X^ε, as it 's the case in [19]. However, this scheme remains to be further investigated.

The two numerical schemes above are not directly implemented in practice and require an important procedure to simulate the conditional expectation. However, there exist different technics which can be adapted to our setting to compute this conditional expectation and we shall only mention the papers: [3], [6] , [9] and [18].

The paper is organized as follows. In Section 2, we prove the convergence of the approximated scheme. In Section 3, we describe discrete-time scheme introduced in [5] and state our first main convergence result. In section 4, we extend the new discrete scheme of [19] and state our second main result. We also discuss a general case of BSDE. Section 5, is devoted to Malliavin calculus for a class of FBSDE with jumps, we then get the L^p-Hölder continuity of Z^ε and Γ^ε via the trace of the Malliavin derivatives of Y^ε.

2 Approximation of decoupled FBSDE driven by pure jump Lévy processes

Let $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F})_{t\leq T}, \mathbb{P})$ be a stochastic basis such that $\mathcal{F}_0$ contains the $\mathbb{P}$-null sets, $\mathcal{F}_T = \mathcal{F}$ and $\mathbb{F}$ satisfies the usual assumptions. We assume that $\mathbb{F}$ is generated by a one-dimensional Brownian motion W and an independent Poisson measure μ on $[0,T] \times E$. We denote by $\mathbb{F}^W = (\mathcal{F}^W_t)_{t\leq T}$ (resp. $\mathbb{F}^\mu = (\mathcal{F}^\mu_t)_{t\leq T}$) the $\mathbb{P}$-augmentation of the natural filtration of W (resp. μ). As usual, we denote by $\mathcal{B}(X)$ the Borel set of topological set X. We introduce the following subset: $E^\varepsilon := \{e \in \mathbb{R},\ \text{s.t} \ / \ |e| \leq \varepsilon\}$, $E_\varepsilon := \{e \in \mathbb{R},\ \text{s.t} \ / \ |e| > \varepsilon\}$, $E := \mathbb{R} = E^\varepsilon \cup E_\varepsilon$.

The martingale measure $\bar{\mu}$ is the compensated measure corresponding to Poisson random measure μ, such that $\bar{\mu}(de, dr) = \mu(de, dr) - \nu(de)dr$, where ν is a Lévy measure on E endowed with its Borel tribe $\mathcal{E}$. The Lévy measure ν will be assumed to satisfy $\nu(\mathbb{R}) = \infty$ and $\int_{\mathbb{R}} |e|^2 \nu(de) < \infty$. Throughout this paper we deal with the measure $\bar{M}$ defined by

$$\bar{M}(t, B) = \int_{[0,t]\times B} e\bar{\mu}(dr, de), \qquad B \in \mathcal{B}(E)$$

which can be considered as a compensated Poisson random measure on $[0,T] \times E$ and $\int_{[0,t]\times E} e\mu(dr, de)$ is a compound Poisson random variable. We associate to $\bar{M}$ the σ-finite measure

$$m(B) := \int_B e^2 \nu(de) \qquad B \in \mathcal{B}(E). \tag{2.1}$$

In particular, we have $\sigma(\varepsilon)^2 = m(E^\varepsilon)$.

The measure $\bar{M}$ is taken to drive the jump noise instead of $\bar{\mu}$, in the aim to adopt the concept of Malliavin calculus on the canonical Lévy space from [12].

For some constant $K > 0$, we consider four K-Lipschitz functions with bounded derivatives $\beta : \mathbb{R} \to \mathbb{R}$, $b : \mathbb{R} \to \mathbb{R}$, $g : \mathbb{R} \to \mathbb{R}$ and $f : \Omega \times \mathbb{R} \times \mathbb{R} \times L^2(E, \mathcal{E}, \nu, \mathbb{R}) \to \mathbb{R}$, where the first

derivative of b, β and g are bounded.

Define ρ to be a measurable function $\rho : E \to \mathbb{R}$ such that:

$$\sup_{e\in E} |\rho(e)| < K. \tag{2.2}$$

For any $p \geq 2$ we consider the following class of processes:

- S^p is the set of real valued adapted rcll process Y such that:

$$\|Y\|_{S^p} := \mathbb{E}\left(\sup_{0\leq t\leq T} |Y_t|^p\right)^{\frac{1}{p}} < \infty.$$

- H^p is the set of progressively measurable $\mathbb{R}$-valued processes Z such that:

$$\|Z\|_{H^p} := \left(\mathbb{E}\left(\int_0^T |Z_r|^2 dr\right)^{\frac{p}{2}}\right)^{\frac{1}{p}} < \infty.$$

- L^p is the set of $\mathcal{P}\otimes\mathcal{E}$ measurable map $U : \Omega\times[0,T]\times E \to \mathbb{R}$ such that:

$$\|U\|_{L^p} := \left(\mathbb{E}\int_0^T\int_E |U_r(e)|^p \nu(de)dr\right)^{\frac{1}{p}} < \infty.$$

- The space $\mathcal{B}^p := \mathcal{S}^p \times H^p \times L^p$ is endowed with the norm

$$\|(Y,Z,U)\|_{\mathcal{B}^p} := \left(\|Y\|^p_{S^p} + \|Z\|^p_{H^p} + \|U\|^p_{L^p}\right)^{\frac{1}{p}}.$$

- $M^{2,p}$ the class of square integrable random variable F of the form:

$$F = \mathbb{E}F + \int_0^T U_r dW_r + \int_0^T\int_E \psi(r,e)\bar{\mu}(de,dr),$$

  where u (resp. ψ) is a progressively measurable (resp. measurable) process satisfying $\sup_{t\leq T}\mathbb{E}|u_t|^p < \infty$ (resp. $\sup_{t\leq T}\mathbb{E}\int_E |\psi(t,e)|^p\nu(de) < \infty$).

2.1 Approximation scheme

In this subsection, we show that the approximation error

$$\begin{aligned} Err^2_\varepsilon(Y,V) \;:=\; & \mathbb{E}\left[\sup_{t\leq T}|Y_t - Y^\varepsilon_t|^2\right] + \mathbb{E}\left[\sup_{t\leq T}\left|\int_0^t V_r dR_r - \int_0^t Z^\varepsilon_r dW_r\right|^2\right] \\ & + \mathbb{E}\left[\int_0^T\int_{E_\varepsilon} |V_r - U^\varepsilon_r(e)|^2 m(de)dr\right], \end{aligned}$$

converges to 0 as ε goes to 0.

Theorem 2.1 *Under the space* $(\Omega, \mathcal{F}, \mathbb{P})$,

1. *There exist a solution* X *on* $[0,T]$ *of*

$$X_t = X_0 + \int_0^t b(X_r)dr + \int_0^t \beta(X_{r^-})dL_r, \tag{2.3}$$

where $X_0 \in \mathbb{R}$.

2. *There exist a solution* X^ε *on* $[0,T]$ *of*

$$X_t^\varepsilon = X_0^\varepsilon + \int_0^t b(X_r^\varepsilon)dr + \int_0^t \beta(X_r^\varepsilon)\sigma(\varepsilon)dW_r + \int_0^t \int_{E_\varepsilon} \beta(X_{r^-}^\varepsilon)\bar{M}(dr,de), \tag{2.4}$$

where $X_0^\varepsilon \in \mathbb{R}$.

Moreover

$$\mathbb{E}\left(\sup_{0\le t\le T} |X_t - X_t^\varepsilon|^2\right) \le C\sigma(\varepsilon)^2. \tag{2.5}$$

For the proof, we state firstly the following Lemma

Lemma 2.1 *On the space* $(\Omega, \mathcal{F}, \mathbb{P})$, *fixing* $\varepsilon > 0$, *we have for* $p \ge 2$:

$$\mathbb{E}\left(\sup_{0\le t\le T} |X_t|^p\right) < \infty. \tag{2.6}$$

$$\mathbb{E}\left(\sup_{0\le t\le T} |X_t^\varepsilon|^p\right) < \infty. \tag{2.7}$$

Proof. We denote by C a constant whose value may change from line to line. Using Jensen's inequality, Burkholder-Davis-Gundy inequality and Lipschitz property of b and β we have:

$$\begin{aligned}
\mathbb{E}\sup_{0\le s\le t}|X_s|^p &\le C\sup_{0\le s\le t}\left(\mathbb{E}\left[|X_0|^p + \left(\int_0^s b(X_r)dr\right)^p + \int_0^s\int_E \beta(X_r)\bar{M}(dr,de)\right]^p\right)\\
&\le C\left(|X_0|^p + \int_0^t \mathbb{E}\Big[|b(X_0)| + |X_r| + |X_0|\Big]^p dr\right.\\
&\quad \left. + \int_0^t\int_E \mathbb{E}\Big[|\beta(X_0)| + |X_0| + |X_r|\Big]^p e^p\nu(de)dr\right)\\
&\le C\left(|X_0|^p + |b(X_0)|^p + |\beta(X_0)|^p + \int_0^t \mathbb{E}\Big[\sup_{0\le u\le r}|X_u|^p dr\Big]\right),
\end{aligned}$$

where C depends on t, $b(X_0)$ and $\beta(X_0)$. We conclude the first assertion by Gronwall's Lemma.

Following the same arguments, we obtain the second assertion. □

Proof of Theorem 2.1 The existence and uniqueness of such SDEs was already studied in the literature see e.g. [17] and [1]. Then it remains to prove the estimate (2.5).

Using Jensen's inequality leads to :

$$
\begin{aligned}
\mathbb{E} \sup_{0\le u\le t} |X_u - X_u^\varepsilon|^2 &\le C\Bigg(\mathbb{E}\left[\int_0^t |b(X_r) - b(X_r^\varepsilon)| dr\right]^2 \\
&+ \mathbb{E}\left[\int_0^t \int_{E_\varepsilon} |\beta(X_r) - \beta(X_r^\varepsilon)| \bar{M}(dr, de)\right]^2 \\
&+ \mathbb{E}\left[\left(\int_0^t \int_{E^\varepsilon} |\beta(X_r)| \bar{M}(de, dr)\right)^2 + \left(\int_0^t |\beta(X_r^\varepsilon)\sigma(\varepsilon)| dW_r\right)^2\right]\Bigg).
\end{aligned}
$$

By Burkholder-Davis-Gundy inequality, we get

$$
\begin{aligned}
\mathbb{E} \sup_{0\le u\le t} |X_u - X_u^\varepsilon|^2 &\le C\Bigg(t\mathbb{E}\left[\int_0^t \left(b(X_r) - b(X_r^\varepsilon)\right)^2 dr\right] \\
&+ \mathbb{E}\left[\int_0^t \int_{E_\varepsilon} \left(\beta(X_r) - \beta(X_r^\varepsilon)\right)^2 m(de)dr\right] \\
&+ \mathbb{E}\left[\int_0^t \int_{E^\varepsilon} \beta(X_r)^2 m(de)dr\right] + C\mathbb{E}\left[\int_0^t \beta(X_r^\varepsilon)^2\sigma(\varepsilon)^2 dr\right]\Bigg).
\end{aligned}
$$

By Lipschitz property of b and β

$$
\begin{aligned}
\mathbb{E} \sup_{0\le u\le t} |X_u - X_u^\varepsilon|^2 &\le C\Bigg(\left[\int_0^t \mathbb{E}\left(X_r - X_r^\varepsilon\right)^2 dr\right] \\
&+ \sigma(\varepsilon)^2 \mathbb{E}\int_0^t \left[\beta^2(X_0) + \beta^2(X_0^\varepsilon) + |X_r^\varepsilon|^2 + |X_r|^2\right] dr\Bigg) \\
&\le C\left[\int_0^t \mathbb{E}\left[\sup_{u\le r} |\delta X_u|^2\right] dr + \sigma(\varepsilon)^2\right].
\end{aligned}
$$

The result follows from Gronwall's Lemma. □

Finally, we can now state the main result of this section.

Theorem 2.2 *Under the space* $(\Omega, \mathcal{F}, \mathbb{P})$,

1. *There exist a unique pair* $(Y, V) \in \mathcal{S}^2 \times H^2$, *which solves the BSDE:*

$$
Y_t = g(X_T) + \int_t^T f(\Theta_r)dr - \int_t^T V_r dL_r, \qquad (2.8)
$$

   where $\Theta := \left(X, Y, \int_E \rho(e) V e\nu(de)\right)$.

2. *For a fixed* $\varepsilon > 0$, *There exist a unique solution* $(Y^\varepsilon, Z^\varepsilon, U^\varepsilon) \in \mathcal{B}^2$ *of the following BSDE:*

$$
Y_t^\varepsilon = g(X_T^\varepsilon) + \int_t^T f(\Theta_r^\varepsilon)dr - \int_t^T Z_r^\varepsilon dW_s - \int_t^T \int_{E_\varepsilon} U_r^\varepsilon(e)\bar{M}(dr, de). \quad (2.9)
$$

   with $\Theta^\varepsilon := \left(X^\varepsilon, Y^\varepsilon, \Gamma^\varepsilon\right)$ *and* $\Gamma^\varepsilon := \int_{E_\varepsilon} \rho(e) U^\varepsilon(e) e\nu(de)$.

Moreover, if $\sup_{t\leq T}\mathbb{E}|V_t|^2<\infty$, *then there exist a constant* C *such that:*

$$Err^2_\varepsilon(Y,V)\leq C\sigma(\varepsilon)^2. \tag{2.10}$$

Remark 2.1 *Observe that*

$$\begin{aligned}\mathbb{E}\left[\sup_{t\leq T}|Y_t-Y^\varepsilon_t|^2\right] \;&+\; \mathbb{E}\left[\sup_{t\leq T}\left|\int_t^T\int_{\mathbb{R}}V_r\bar{M}(de,dr)\right.\right.\\ &\qquad\left.\left.-\int_0^t Z^\varepsilon_r dW_r-\int_0^T\int_{E_\varepsilon}U^\varepsilon_r(e)\bar{M}(de,dr)\right|^2\right]\\ &\leq\; Err^2_\varepsilon(Y,V)\\ &\leq\; \sigma(\varepsilon)^2.\end{aligned}$$

Which shows clearly the convergence of the approximated scheme (1.3) to (1.1).

Proof. of Theorem 2.2 Existence and uniqueness of the solutions of BSDEs (2.8) and (2.9) was already proved, see e.g [4].

We are going to prove inequality (2.10). By Itô's formula applied to $|\delta Y|^2:=|Y-Y^\varepsilon|^2$ yields :

$$\begin{aligned}|\delta Y_t|^2 \;&+\; \int_t^T Z^{\varepsilon 2}_r dr+\int_t^T\int_{E_\varepsilon}\left(U^\varepsilon_r(e)-V_r\right)^2 m(de)dr\\ &=\; |g(X_T)-g(X^\varepsilon_T)|^2+\sigma(\varepsilon)^2\int_t^T V_r^2dr-\int_t^T\delta Y_r Z^\varepsilon_r dW_r\\ &\quad +2\int_t^T\delta Y_r\left(f\left(\Theta_r\right)-f\left(\Theta^\varepsilon_r\right)\right)dr+2\int_t^T\int_{E^\varepsilon}\left[(\delta Y_{s^-}+V_r)^2-\delta Y^2_{s^-}\right]\bar{M}(de,dr)\\ &\quad -2\int_t^T\int_{E_\varepsilon}\left[(\delta Y_{s^-}+U^\varepsilon_r(e)-V_r)^2-\delta Y^2_{s^-}\right]\bar{M}(de,dr).\end{aligned} \tag{2.11}$$

Taking expectation in both hand-side of the above equality we get

$$\begin{aligned}\mathbb{E}\;&\left[\delta Y_t^2+\int_t^T Z^{\varepsilon 2}_r dr+\int_t^T\int_{E_\varepsilon}\left(U^\varepsilon_r(e)-V_r\right)^2 m(de)dr\right]\\ &=\mathbb{E}\left[|g(X_T)-g(X^\varepsilon_T)|^2+\sigma(\varepsilon)^2\int_t^T V_r^2dr+2\int_t^T\delta Y_r\left(f\left(\Theta_r\right)-f\left(\Theta^\varepsilon_r\right)\right)dr\right].\end{aligned}$$

From Lemma 2.1, Lipschitz property of g and Jensen inequality we obtain:

$$\begin{aligned}\mathbb{E}\;&\left[|Y_t-Y^\varepsilon_t|^2+\int_t^T Z^{\varepsilon}{}^2_r dr+\int_t^T\int_{E_\varepsilon}\left(U^\varepsilon_r(e)-V_r\right)^2 m(de)dr\right]\\ \leq\;& C\left(\sigma(\varepsilon)^2+K\mathbb{E}\int_t^T(Y_r-Y^\varepsilon_r)(|X_r-X^\varepsilon_r|+|Y_r-Y^\varepsilon_r|)dr\right.\\ &+K\mathbb{E}\int_t^T\left[(Y_r-Y^\varepsilon_r)\int_{E_\varepsilon}\rho(e)e|U^\varepsilon_r(e)-V_r|\nu(de)\right]dr\\ &\left.+\,K\mathbb{E}\int_t^T\left[(Y_r-Y^\varepsilon_r)\int_{E^\varepsilon}\rho(e)V_r e\nu(de)\right]dr\right).\end{aligned}$$

Using the fact that $ab \leq \alpha a^2 + \frac{1}{\alpha} b^2$ for some $\alpha > 0$, yields to

$$\begin{aligned}
\mathbb{E} & \left[|Y_t - Y_t^\varepsilon|^2 + \int_t^T Z^{\varepsilon 2}_r dr + \int_t^T \int_{E_\varepsilon} (U_r^\varepsilon(e) - V_r)^2 \, m(de)dr \right] \\
\leq & \; C \Bigg(\sigma(\varepsilon)^2 + K(1+\alpha^2+\gamma^2+\eta^2) \mathbb{E} \int_t^T (Y_r - Y_r^\varepsilon)^2 dr + \frac{K}{\alpha^2} \mathbb{E} \int_t^T |X_r - X_r^\varepsilon|^2 dr \\
& + \frac{K}{\gamma^2} \mathbb{E} \int_t^T \int_{E_\varepsilon} \rho(e)^2 |U_r^\varepsilon(e) - V_r|^2 m(de)dr + \frac{K}{\eta^2} \mathbb{E} \int_t^T \int_{0 \leq |e| \leq \varepsilon} \rho(e)^2 V_r^2 m(de)dr \Bigg),
\end{aligned}$$

where α and γ are two constants taken such that $\frac{K}{\alpha^2} = \frac{K^3}{\gamma^2} = \frac{1}{2}$, we then get

$$\begin{aligned}
\mathbb{E}\Big[|Y_t - Y_t^\varepsilon|^2 & + \int_t^T Z^{\varepsilon 2}_r dr + \int_t^T \int_{E_\varepsilon} (U_r^\varepsilon(e) - V_r)^2 \, m(de)dr \Big] \\
& \leq C \left(\sigma(\varepsilon)^2 + (K + 2K^2 + 2KK^2) \mathbb{E} \int_t^T (Y_r - Y_r^\varepsilon)^2 dr \right). \qquad (2.12)
\end{aligned}$$

Using Gronwall's Lemma, we deduce that

$$\mathbb{E}|Y_t - Y_t^\varepsilon|^2 \leq C\sigma(\varepsilon)^2. \qquad (2.13)$$

Plugging this estimate in the previous upper bound, we get

$$\mathbb{E} \int_0^T Z_r^{\varepsilon 2} dr + \mathbb{E} \int_0^T \int_{E_\varepsilon} (U_r^\varepsilon(e) - V_r)^2 \, m(de)dr \leq C\sigma(\varepsilon)^2, \qquad (2.14)$$

Then

$$\mathbb{E}|\delta Y_t|^2 + \mathbb{E} \int_0^T Z_r^{\varepsilon 2} dr + \mathbb{E} \int_0^T \int_{E_\varepsilon} (U_r^\varepsilon(e) - V_r)^2 \, m(de)dr \leq C\sigma(\varepsilon)^2. \qquad (2.15)$$

Now using Burkholder-Davis-Gundy inequality, we have

$$\mathbb{E} \sup_{t \leq T} |\delta Y_t|^2 + \mathbb{E} \int_0^T Z_r^{\varepsilon 2} dr + \mathbb{E} \int_0^T \int_{E_\varepsilon} (U_r^\varepsilon(e) - V_r)^2 \, m(de)dr \leq C\sigma(\varepsilon)^2. \qquad (2.16)$$

From other side, it follows by Burkholder-Davis-Gundy inequality and (2.16) that:

$$\begin{aligned}
\mathbb{E} \left[\sup_{t \leq T} \left| \int_0^t Z_r^\varepsilon dW_r - \int_0^t V_r dR_r \right|^2 \right] & \leq C\mathbb{E} \left[\sup_{t \leq T} \left| \int_0^t Z_r^\varepsilon dW_r \right|^2 + \sup_{t \leq T} \left| \int_0^t V_r dR_r \right|^2 \right] \\
& \leq C\mathbb{E} \left[\int_0^T Z_r^{\varepsilon 2} dr + \sigma(\varepsilon^2) \int_0^T V_r^2 dr \right] \\
& \leq \mathbb{E} \int_0^T Z_r^{\varepsilon 2} dr + C\sigma(\varepsilon^2). \qquad (2.17)
\end{aligned}$$

The result now follows by combining (2.16) and (2.17). □

Remark 2.2 *One can show that* $\sup_{t\leq T}\mathbb{E}|V_t|^2<\infty$*, following the same arguments as in Remark 2.8 in [19].*

Remark 2.3 *In the BSDE (2.9), each time we change* ε*, there exist a unique pair* $(Z^\varepsilon, U^\varepsilon)$ *of predictable process, such that the BSDE (2.9) has a solution.*

3 Forward-backward Euler scheme

In this section, we discretize the solution $(X^\varepsilon, Y^\varepsilon, Z^\varepsilon, \Gamma^\varepsilon)$ of (1.3) by $(X^\pi, Y^\pi, Z^\pi, \Gamma^\pi)$ defined by induction in (3.3) and then we show the convergence of $(X^\pi, Y^\pi, Z^\pi, \Gamma^\pi)$ to the solution of (1.2). Thus let us recall some definition and notation.
For each $t\in[t_i, t_{i+1})$, we define:

$$\bar{Z}_t = n\mathbb{E}\left[\int_{t_i}^{t_{i+1}} Z_s ds/\mathcal{F}_{t_i}\right], \qquad \bar{\Gamma}_t = n\mathbb{E}\left[\int_{t_i}^{t_{i+1}} \Gamma_s ds/\mathcal{F}_{t_i}\right], \tag{3.1}$$

and

$$\bar{Z}^\pi_{t_i} = n\mathbb{E}\left[\int_{t_i}^{t_{i+1}} Z^\pi_s ds/\mathcal{F}_{t_i}\right], \qquad \bar{\Gamma}^\pi_{t_i} = n\mathbb{E}\left[\int_{t_i}^{t_{i+1}} \Gamma^\pi_s ds/\mathcal{F}_{t_i}\right]. \tag{3.2}$$

The process $\bar{Z}_{t_i}$ and $\bar{\Gamma}_{t_i}$ (resp. $\bar{Z}^\pi_{t_i}$ and $\bar{\Gamma}^\pi_{t_i}$) can be interpreted as the best approximation of Z_{t_i} and Γ_{t_i} (resp. $Z^\pi_{t_i}$ and $\Gamma^\pi_{t_i}$). We know from Bouchard and Elie [5], that FBSDE (2.8) has a backward Euler scheme taking the form:

$$\left\{\begin{array}{lcl} \bar{X}^\pi_{t_i+1} & = & \bar{X}^\pi_{t_i} + \frac{1}{n}b(\bar{X}^\pi_{t_i}) + \sigma(\varepsilon)\Delta W_{i+1} + \int_{E_\varepsilon}\beta(\bar{X}^\pi_{t_i})\bar{M}(de,(t_i,t_{i+1}]) \\ \bar{Z}^\pi_t & = & n\mathbb{E}\left[\bar{Y}^\pi_{t_{i+1}}\Delta W_{i+1}/\mathcal{F}_{t_i}\right] \\ \bar{\Gamma}^\pi_t & = & n\mathbb{E}\left[\bar{Y}^\pi_{t_{i+1}}\int_{E_\varepsilon}\rho(e)\bar{M}(de,(t_i,t_{i+1}])/\mathcal{F}_{t_i}\right] \\ \bar{Y}^\pi_t & = & \mathbb{E}\left[\bar{Y}^\pi_{t_{i+1}}/\mathcal{F}_{t_i}\right] + \frac{1}{n}f(\bar{X}^\pi_{t_i}, \bar{Y}^\pi_{t_i}, \bar{\Gamma}^\pi_{t_i}) \end{array}\right. \tag{3.3}$$

for which the discretization error:

$$\begin{aligned} \overline{Err}_n(Y^\varepsilon, Z^\varepsilon, \Gamma^\varepsilon) &:= \left\{\sup_{t\leq T}\mathbb{E}\left[|Y^\varepsilon_t - \bar{Y}^\pi_t|^2\right] + \|Z^\varepsilon - \bar{Z}^\pi\|^2_{H^2} + \|\Gamma^\varepsilon - \bar{\Gamma}^\pi\|^2_{H^2}\right\}^{\frac{1}{2}} \\ &\leq Cn^{-1/2}, \end{aligned} \tag{3.4}$$

converges to 0 as the discretization step $\frac{T}{n}$ tends to 0. Means that the discretization scheme (3.3) achieves the optimal convergence rate $n^{-1/2}$. The regularity of Z^ε and Γ^ε has been studied in L^2 sense in [5] when the terminal value is a functional of forward diffusion.

It is well known also that

$$\max_{i<n}\mathbb{E}\left[\sup_{t\in[t_i,t_{i+1}]}|X^\varepsilon_{t_i} - \bar{X}^\pi_{t_i}|^2\right] \leq Cn^{-1/2}.$$

Our aim in this part, is to show that the approximation-discretization error between BSDE (1.2) and (3.3):

$$\begin{aligned}\overline{Err}^2_{(n,\varepsilon)}(Y,V) &:= \sup_{t\le T}\mathbb{E}\left[|Y_t-\bar{Y}^\pi_t|^2\right]+\sup_{t\le T}\mathbb{E}\left|\int_0^t V_r dR_r-\int_0^t \bar{Z}^\pi_r dW_r\right|^2\\ &\quad +\|\Gamma-\bar{\Gamma}^\pi\|^2_{H^2},\end{aligned}\tag{3.5}$$

converges to 0 as $(\varepsilon,n)\to(0,\infty)$.

The first main result of this paper is:

Proposition 3.1 *Assuming Lipschitz property of coefficients b and β, the approximation-discretization error defined in (3.5) is bounded by:*

$$\overline{Err}_{(n,\varepsilon)}(Y,V)\le C\left(n^{-1/2}+\sigma(\varepsilon)\right).\tag{3.6}$$

Means that:

$$\overline{Err}_{(n,\varepsilon)}(Y,V)\underset{(n,\varepsilon)\to(\infty,0)}{\longrightarrow} 0.$$

Proof. From (3.5), Jensen inequality and Burkholder-Davis-Gundy inequality, we have:

$$\begin{aligned}\overline{Err}^2_{(n,\varepsilon)}(Y,V) &\le C\Big(\sup_{t\le T}\mathbb{E}\left[|Y_t-Y^\varepsilon_t|^2+|Y^\varepsilon_t-\bar{Y}^\pi_t|^2\right]+\|\Gamma-\Gamma^\varepsilon\|^2_{H^2}+\|\Gamma^\varepsilon-\bar{\Gamma}^\pi\|^2_{H^2}\\ &\quad+\sup_{t\le T}\mathbb{E}\left|\int_0^t V_r dR_r-\int_0^t Z^\varepsilon_r dW_r\right|^2+\|Z^\varepsilon-\bar{Z}^\pi\|^2_{H^2}\Big).\end{aligned}\tag{3.7}$$

From other side, it follows from Hölder inequality that:

$$\int_{E_\varepsilon}\rho(e)e(V_r-U^\varepsilon_r(e))\nu(de)\le K\left(\int_{E_\varepsilon}(V_r-U^\varepsilon_r(e))^2 m(de)\right)^{\frac{1}{2}}\nu\Big(E_\varepsilon\Big)^{\frac{1}{2}}.\tag{3.8}$$

Recalling that $\nu(E_\varepsilon)<\infty$ so that ν has a.s. only a finite number of big jumps on $[0,T]$. Combining the two last inequalities with (3.4) leads to:

$$\begin{aligned}\overline{Err}^2_{(n,\varepsilon)}(Y,V) &\le C\Bigg(n^{-1}+\sup_{t\le T}\mathbb{E}\left[|Y_t-Y^\varepsilon_t|^2\right]+\sup_{t\le T}\mathbb{E}\left|\int_0^t V_r dR_r-\int_0^t Z_r dW_r\right|^2\\ &\quad+\int_0^T\int_{E^\varepsilon}\mathbb{E}|V_r|^2 m(de)dr+\mathbb{E}\left[\int_0^T\int_{E_\varepsilon}|U^\varepsilon_r(e)-V_r|^2 m(de)dr\right]\Bigg).\end{aligned}$$

By Theorem 2.2 we get:

$$\overline{Err}_{(n,\varepsilon)}(Y,V)\ \le\ C\left(n^{-1/2}+\sigma(\varepsilon)\right),\tag{3.9}$$

where C depends on K. □

Remark 3.1 *In the general case, as we neglect the small jump, the Brownian part in (1.3) disappears. In this case the assertion (3.6) can be replaced by:*

$$\overline{Err}_{(n,\varepsilon)}(Y,V)\le Cn^{-1/2}.$$

Remark 3.2 *Taking $\varepsilon=n^{-1/2}$, we obtain the optimal convergence rate $n^{-1/2}$ in (3.6),*

$$\overline{Err}_{(n,\varepsilon)}(Y,V)\le Cn^{-1/2},$$

which is exactly the approximation error in [5].

4 A discrete scheme via Malliavin derivatives

In this section, we generalize the new discrete scheme recently introduced by Hu, Nualard and Song [19] from a general BSDE, to our framework of decoupled forward-backward SDEs with jumps. For this aim, we use the Malliavin derivatives of Y to derive the discrete scheme. We first fix a regular grid $\pi := \{t_i := iT/n, i = 0, ..., n\}$ on $[0, T]$ and approximate the forward SDE X^ε in (1.3) by its Euler scheme X^π already defined in (3.3). It'is hard to prove existence and convergence of Malliavin derivatives of $\bar{X}^\pi$. However, to avoid this problem, we can instead consider the continuous-time version of the Euler scheme, then we define the function ϕ for each $t \in [0, T]$:

$$\phi_t^n := \max\{t_i,\ i = 0, ..., n.\ /\ t_i \le t\}, \tag{4.1}$$

for which we associate:

$$X_t^\pi \ := \ X_{\phi_t^n}^\pi + \sigma(\varepsilon) b(X_{\phi_t^n}^\pi)(t - \phi_t^n) + \sigma(\varepsilon)\beta(X_{\phi_t^n}^\pi)(W_t - W_{\phi_t^n}) + \int_{E_\varepsilon} \beta(X_{\phi_t^n}^\pi)\bar{M}(de, (t, \phi_t^n]).$$

It could be written as

$$X_t^\pi \ := \ X_0 + \int_0^t b(X_{\phi_r^n}^\pi) dr + \int_0^t \sigma(\varepsilon)\beta(X_{\phi_r^n}^\pi) dW_r + \int_0^t \int_{E_\varepsilon} \beta(X_{\phi_r^n}^\pi)\bar{M}(de, dr).$$

It is well known that under Lipschitz property of the coefficients

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_t^\varepsilon - X_t^\pi|^p\right]^{1/p} \le C\pi^{1/2}. \tag{4.2}$$

The Malliavin derivatives of the continuous-time version of Euler scheme for $\theta \le s$ a.e. are:

$$\begin{aligned} D_\theta X_t^\pi \ &= \ \int_\theta^t \partial_x b(X_{\phi_r^n}^\pi) D_\theta X_{\phi_r^n}^\pi dr + \int_\theta^t \int_{E_\varepsilon} \partial_x \beta(X_{\phi_r^n}^\pi) D_\theta X_{\phi_r^n}^\pi \bar{M}(de, dr) + \sigma(\varepsilon)\beta(X_\theta^\pi) \\ &\quad + \sigma(\varepsilon) \int_\theta^t \partial_x \beta(X_{\phi_r^n}^\pi) D_\theta X_{\phi_r^n}^\pi dW_r, \\ D_{\theta,e} X_t^\pi \ &= \ \int_\theta^t D_{\theta,e} b(X_r^\pi) dr + \int_\theta^t \int_{E_\varepsilon} D_{\theta,e}\beta(X_r^\pi)\bar{M}(de, dr) + \sigma(\varepsilon)\int_\theta^t D_{\theta,e}\beta(X_r^\pi) dW_r \\ &\quad + \beta(X_\theta^\pi). \end{aligned}$$

We introduce some additional assumptions:

(A1) $f(t, y, \gamma)$ doesn't depend on x.

(A2) The first derivative of b and β and g is a K-Lipschitz function

$$|b'(x) - b'(y)| + |\beta'(x) - \beta'(y)| + |g'(x) - g'(y)| \le K|x - y|.$$

(A3) $f(t, y, \gamma)$ is linear with respect to t, y and γ. Moreover, there exist three bounded functions f_1, f_2 and f_3 such that :

$$f(t, y, u) = f_1(t) + f_2(t) y + f_3(t)\gamma. \tag{4.3}$$

Lemma 4.1 *Under Lipschitz continuity of b and β, we have for any $q \geq 1$*

$$\sup_{0\leq\theta\leq T}\sup_{n\geq 1}\mathbb{E}\left[\sup_{\theta\leq t\leq T}\|D_\theta X_t^\pi\|^{2q}\right] < \infty, \quad (4.4)$$

$$\sup_{0\leq\theta\leq T}\sup_{n\geq 1}\mathbb{E}\left[\sup_{\theta\leq t\leq T}\|D_{\theta,e} X_t^\pi\|^{2q}\right] < \infty. \quad (4.5)$$

For the proof see the Appendix.

We then derive the following theorem

Theorem 4.1 *Under assumption (A2), Lipschitz continuity of b and β and for any $p \geq 2$, we have,*

$$\mathbb{E}\left[\sup_{t\in[0,T]}|D_\theta X_t^\varepsilon - D_\theta X_t^\pi|^p\right]^{1/p} \leq C_p \pi^{1/2}. \quad (4.6)$$

$$\mathbb{E}\left[\sup_{t\in[0,T]}|D_{\theta,e} X_t^\varepsilon - D_{\theta,e} X_t^\pi|^p\right]^{1/p} \leq C_p \pi^{1/2}. \quad (4.7)$$

Proof. Using Burkholder-Davis-Gundy inequality, Jensen inequality, inequality (4.2) and Lemma 4.1

$$\begin{aligned}
\mathbb{E}\left[\sup_{s\in[0,t]}|D_\theta X_t^\varepsilon - D_\theta X_t^\pi|^p\right] \leq\ & C_p\mathbb{E}\left[\int_0^t \left|\partial_x b(X^\pi_{\phi_r^n})D_\theta X^\pi_{\phi_r^n} - \partial_x b(X_r)D_\theta X_r^\varepsilon\right|^p dr\right. \\
& + \int_0^t\int_{E_\varepsilon}\left|\partial_x\beta(X^\pi_{\phi_r^n})D_\theta X^\pi_{\phi_r^n} - \partial_x\beta(X_r^\varepsilon)D_\theta X_r^\varepsilon\right|^p \nu(de)dr \\
& + \sigma^p(\varepsilon)\left(\int_0^t\left|\partial_x\beta(X^\pi_{\phi_r^n})D_\theta X^\pi_{\phi_r^n} - \partial_x\beta(X_r^\varepsilon)D_\theta X_r^\varepsilon\right|^2 dr\right)^{p/2} \\
& \left. + \sigma^p(\varepsilon)\left|\beta(X_\theta^\pi) - \beta(X_\theta^\varepsilon)\right|^p\right].
\end{aligned}$$

which leads to

$$\begin{aligned}
\leq\ & C_p\mathbb{E}\left[\int_0^t \left|D_\theta X_r^\varepsilon\right|^p\left[|\partial_x b(X_r^\varepsilon) - \partial_x b(X^\pi_{\phi_r^n})|^p + |\partial_x\beta(X_r^\varepsilon) - \partial_x\beta(X^\pi_{\phi_r^n})|^p\right]dr\right. \\
& + \int_0^t\left|D_\theta X_r^\varepsilon - D_\theta X^\pi_{\phi_r^n}\right|^p\left[|\partial_x b(X^\pi_{\phi_r^n})|^p + |\partial_x\beta(X^\pi_{\phi_r^n})|^p\right]dr \\
& \left. + |X_\theta^\pi - X_\theta^\varepsilon|^p\right] \\
\leq\ & C_p\mathbb{E}\left(\pi^{p/2} + \int_0^t \sup_{u\in[0,r]}|D_\theta X_u^\varepsilon - D_\theta X_u^\pi|^p dr\right).
\end{aligned}$$

We conclude by using Gronwall's Lemma. Following the same arguments, we prove the second assertion.

□

Now we derive the discrete scheme using the expression of Z^ε and U^ε as the trace of

Malliavin derivatives of Y. From equation (5.15) and (5.17), the two Malliavin derivatives $D_\theta Y_t$, $D_{\theta,e} Y_t$ could be expressed as:

$$
\begin{aligned}
D_\theta Y_t^\varepsilon &= \mathbb{E}\left(\mathcal{E}_{t,T}\partial_x g(X_T^\varepsilon)D_\theta X_T^\varepsilon + \int_t^T \mathcal{E}_{t,r}\partial_x f(\Theta_r^\varepsilon)D_\theta X_r^\varepsilon dr/\mathcal{F}_t\right) && (4.8)\\
D_{\theta,e} Y_t^\varepsilon &= \mathbb{E}\left(\mathcal{E}_{t,T}^e D_{\theta,e} g(X_T^\varepsilon) + \int_t^T \mathcal{E}_{t,r}^e \alpha_{\theta,r} D_{\theta,e} X_r^\varepsilon dr/\mathcal{F}_t\right), && (4.9)
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{E}_{t,r} &= exp\left\{\int_t^r \left(\partial_y f(\Theta_u^\varepsilon) - \frac{1}{2}\int_{E_\varepsilon} \partial_\gamma f^2(\Theta_u^\varepsilon)\rho^2(e)m(de)\right)du \right.\\
&\qquad \left. + \int_t^r \int_{E_\varepsilon} \partial_\gamma f(\Theta_u^\varepsilon)\rho(e)\bar{M}(de,du)\right\}\\
\mathcal{E}_{t,r}^e &:= exp\left\{\int_t^r \left[\alpha_{\theta,u} - \frac{1}{2}\alpha_{\theta,u}^2 \int_{E_\varepsilon} \rho^2(e)m(de)\right]du + \int_t^r \int_{E_\varepsilon} \alpha_{\theta,u}\rho(e)\bar{M}(de,du)\right\}
\end{aligned}
$$

and

$$
\alpha_{\theta,r} := \frac{f(\Theta_r^\varepsilon + D_{\theta,e}\Theta_r^\varepsilon) - f(\Theta_r^\varepsilon)}{D_{\theta,e}X_r^\varepsilon + D_{\theta,e}Y_r + D_{\theta,e}\Gamma_r} 1_{\{D_{\theta,e}X_r^\varepsilon + D_{\theta,e}Y_r + D_{\theta,e}\Gamma_r \neq 0\}}.
$$

Thus, we define our discrete scheme for $i = n-1, ..., 1, 0.$ and $t \in [t_i, t_{i+1})$ by induction

$$
\left\{
\begin{aligned}
Y_{t_i}^\pi &= \mathbb{E}\Big[Y_{t_{i+1}}^\pi + f(\Theta_{t_{i+1}}^\pi)\Delta t_{i+1}/\mathcal{F}_{t_i}\Big]\\
Z_{t_i}^\pi &= \mathbb{E}\Big[\mathcal{E}_{t_{i+1},t_n}^\pi \partial_x g(X_T^\pi)D_{t_i}X_T^\pi + \textstyle\sum_{k=i}^{n-1} \mathcal{E}_{t_{i+1},t_{k+1}}^\pi \partial_x f(\Theta_{t_{k+1}}^\pi)D_{t_i}X_{t_{k+1}}^\pi \Delta t_k/\mathcal{F}_{t_i}\Big]\\
\Gamma_{t_i}^\pi &= \mathbb{E}\Big[\textstyle\int_{E_\varepsilon} \rho(e)\big[\mathcal{E}_{t_{i+1},t_n}^{e,\pi} D_{t_i,e} g(X_T^\pi) + \sum_{k=i}^{n-1} \mathcal{E}_{t_{i+1},t_{k+1}}^{e,\pi}\alpha_{t_i,t_{k+1}}^\pi D_{t_i,e}X_{t_{k+1}}^\pi \Delta t_k\big]\nu(de)/\mathcal{F}_{t_i}\Big]
\end{aligned}
\right.
\quad (4.10)
$$

with terminal conditions
$Y_{t_n}^\pi = g(X_T^\pi)$, $\quad Z_{t_n}^\pi = \sigma(\varepsilon)\partial_x g(X_T^\pi)\beta(X_T^\pi)$, $\quad U_{t_n,e}^\pi = g(X_T^\pi + \beta(X_T^\pi)) - g(X_T^\pi)$,
where for any $0 \le i < j \le n$,

$$
\begin{aligned}
\mathcal{E}_{t_i,t_j}^\pi &= exp\left\{\sum_{k=i}^{j-1}\int_{t_k}^{t_{k+1}} \left[\partial_y f(\Theta_{t_k}^\pi) - \frac{1}{2}\int_{E_\varepsilon} \partial_\gamma f^2(\Theta_{t_k}^\pi)\rho^2(e)m(de)\right]dr \right.\\
&\qquad \left. + \sum_{k=i}^{j-1}\int_{t_k}^{t_{k+1}} \int_{E_\varepsilon} \partial_\gamma f(\Theta_{t_k}^\pi)\rho(e)\bar{M}(de,dr)\right\}, && (4.11)\\
\mathcal{E}_{t_i,t_j}^{e,\pi} &= exp\left\{\sum_{k=i}^{j-1}\int_{t_k}^{t_{k+1}} \left[\alpha_{\theta,r,t_k}^\pi - \frac{1}{2}{\alpha^\pi}^2_{\theta,r,t_k}\int_{E_\varepsilon} \rho^2(e)m(de)\right]dr \right.\\
&\qquad \left. + \sum_{k=i}^{j-1}\int_{t_k}^{t_{k+1}} \int_{E_\varepsilon} \alpha_{\theta,r,t_k}^\pi \rho(e)\bar{M}(de,dr)\right\}, && (4.12)
\end{aligned}
$$

and

$$
\begin{aligned}
\alpha_{\theta,r,t_k}^\pi &:= \frac{f(\Theta_{t_k}^\pi + D_{\theta,e}\Theta_{t_k}^\pi) - f(\Theta_{t_k}^\pi)}{D_{\theta,e}X_{t_k}^\pi + D_{\theta,e}Y_{t_k}^\pi + D_{\theta,e}\Gamma_{t_k}^\pi} \times 1_{\{D_{\theta,e}X_{t_k}^\pi + D_{\theta,e}Y_{t_k}^\pi + D_{\theta,e}\Gamma_{t_k}^\pi \neq 0\}},\\
\Gamma_{t_k}^\pi &:= \int_{E_\varepsilon} U_{t_k,e}^\pi \rho(e)\nu(de),
\end{aligned}
$$

with $\Theta^\pi_{t_k} = \big(r, X^\pi_{t_k}, Y^\pi_{t_k}, \Gamma^\pi_{t_k}\big)$.

We are going to compute the discretization error of our discrete scheme and prove the convergence of the above scheme. We recall the expression of the error between the solution of (1.3) and (4.10):

$$Err^p_n(Y^\varepsilon, Z^\varepsilon, \Gamma^\varepsilon) \;:=\; \mathbb{E}\max_{0\le i\le n}\left[|Y^\varepsilon_{t_i} - Y^\pi_{t_i}|^p + |Z^\varepsilon_{t_i} - Z^\pi_{t_i}|^p + |\Gamma^\varepsilon_{t_i} - \Gamma^\pi_{t_i}|^p\right],$$

where, $\Gamma^\pi_{t_i} = \int_{E_\varepsilon}\rho(e)U^\pi_{t_i,e}\nu(de)$.

We also recall the expression of discretization-approximation error between (1.2) and (4.10)

$$\begin{aligned}Err^2_{n,\varepsilon}(Y,V) \;:=\; & \max_{0\le i\le n}\sup_{t\in[t_i,t_{i+1}]}\mathbb{E}\left[|Y_t - Y^\pi_{t_i}|^2\right] + \mathbb{E}\left|\int_0^T V_r dR_r - \sum_{i=0}^{n-1} Z^\pi_{t_i}\Delta W_{t_i}\right|^2 \\ & + \sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}\mathbb{E}|\Gamma_t - \Gamma^\pi_{t_i}|^2 dt.\end{aligned}$$

We conclude this section with the following Theorems whose proof are at the end of section 5.

Theorem 4.2 *Under assumption 5.1, we assume the existence of a constant $L_3 > 0$ such that :*

$$|f(t_2,y,u) - f(t_1,y,u)| \le L_3|t_2 - t_1|^{\frac{1}{2}}. \tag{4.13}$$

Then there exist a positive constant C independent of π such that:

$$Err^p_n(Y^\varepsilon, Z^\varepsilon, \Gamma^\varepsilon) \le C_p|\pi|^{\frac{p}{2}-\frac{p}{2log\frac{1}{|\pi|}}}. \tag{4.14}$$

The second main result of this paper is summarized in the following theorem:

Theorem 4.3 *Under the same assumptions as Theorem 4.2, we have*

$$Err_{n,\varepsilon}(Y,V) \le C\left(\sigma(\varepsilon) + |\pi|^{\frac{1}{2}-\frac{1}{2log\frac{1}{|\pi|}}}\right). \tag{4.15}$$

Remark 4.1 *The importance of the above scheme is it can be adapted to a backward SDE when the generator does not depends on the terminal value of a forward equation.*
Consider the following backward stochastic differential equation driven by pure jump Lévy processes

$$\hat{Y}_t \;=\; \hat{\xi} + \int_t^T f\Big(r, \hat{Y}_r, \int_E \rho(e)\hat{V}_r e\nu(de)\Big)dr - \int_t^T \hat{V}_r dL_r \tag{4.16}$$

Which we approximate by

$$\hat{Y}^\varepsilon_t \;=\; \hat{\xi} + \int_t^T f\Big(r, \hat{Y}^\varepsilon_r, \int_{E_\varepsilon}\rho(e)\hat{V}_r e\nu(de)\Big)dr - \int_t^T \hat{Z}^\varepsilon_r dW_s - \int_t^T\int_{E_\varepsilon}\hat{U}^\varepsilon_r(e)\bar{M}(dr,de) \tag{4.17}$$

We finally propose the below discrete time scheme, defined by terminal values $\hat{Y}^\pi_{t_n} = \xi$, $\hat{Z}^\pi_{t_n} = D_T\xi$ *and* $\hat{U}^\pi_{t_i} = D_{T,e}\xi$

$$\left\{\begin{array}{lcl} \hat{Y}^\pi_{t_i} & = & \mathbb{E}\Big[\hat{Y}^\pi_{t_{i+1}} + f(\hat{\Theta}^\pi_{t_{i+1}})\Delta t_{i+1}/\mathcal{F}_{t_i}\Big] \\ \hat{Z}^\pi_{t_i} & = & \mathbb{E}\Big[\mathcal{E}^\pi_{t_{i+1},t_n} D_{t_i}\hat{\xi} + \sum_{k=i}^{n-1}\mathcal{E}^\pi_{t_{i+1},t_{k+1}}\partial_x f(\hat{\Theta}^\pi_{t_{k+1}}) D_{t_i}X^\pi_{t_{k+1}}\Delta t_k/\mathcal{F}_{t_i}\Big] \\ \hat{\Gamma}^\pi_{t_i} & = & \mathbb{E}\Big[\int_{E_\varepsilon}\rho(e)\big[\mathcal{E}^{e,\pi}_{t_{i+1},t_n} D_{t_i,e}\hat{\xi} + \sum_{k=i}^{n-1}\mathcal{E}^{e,\pi}_{t_{i+1},t_{k+1}}\alpha^\pi_{t_i,t_{k+1}} D_{t_i,e}X^\pi_{t_{k+1}}\Delta t_k\big]\nu(de)/\mathcal{F}_{t_i}\Big] \end{array}\right. \tag{4.18}$$

with $\hat{\Theta}^\pi_{t_k} = \Big(r, \hat{Y}^\pi_{t_k}, \hat{\Gamma}^\pi_{t_k}\Big)$.

Under the same assumptions of Theorem 4.2 we prove the convergence of the system (4.18) to BSDE (4.16). Moreover, we obtain the upper bound

$$\max_{0\le i\le n}\sup_{t\in[t_i,t_{i+1}]}\mathbb{E}\left[|Y_t - Y^\pi_{t_i}|^2\right] + \mathbb{E}\left|\int_0^T V_r dR_r - \sum_{i=0}^{n-1} Z^\pi_{t_i}\Delta W_{t_i}\right|^2 + \sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}\mathbb{E}|\Gamma_t - \Gamma^\pi_{t_i}|^2 dt$$
$$\le C\left(\sigma^2(\varepsilon) + |\pi|^{1-\frac{1}{log\frac{1}{|\pi|}}}\right). \tag{4.19}$$

5 Malliavin calculus for FBSDEs

For ease of notations, we shall denote throughout this section the process $(X^\varepsilon, Y^\varepsilon, Z^\varepsilon, \Gamma^\varepsilon)$ by (X, Y, Z, Γ).
In this section, we study some regularity properties of the solution (X, Y, Z, Γ). We recall the system (1.3) using the new notations

$$\left\{\begin{array}{lcl} X_t & = & X_0 + \int_0^t b(X_r)dr + \int_0^t \beta(X_r)\sigma(\varepsilon)dW_r + \int_0^t\int_{E_\varepsilon}\beta(X_{r^-})\bar{M}(dr,de) \\ Y_t & = & g(X_T) + \int_t^T f(\Theta_r)dr - \int_t^T Z_r dW_s - \int_t^T\int_{E_\varepsilon} U_r(e)\bar{M}(dr,de) \end{array}\right. \tag{5.1}$$

In fact, there are many methods to develop Malliavin calculus for Lévy processes. In our paper, we opt for the approach of Solé et al. [25], based on a chaos decomposition in terms of multiple stochastic integrals with respect to the random measure $\bar{M}$. Adopting notation of [12], we will recall the suitable canonical space we adopt to our setting.

We start by introducing some additional notations and definitions. We assume that the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is the product of two canonical spaces $(\Omega_W\times\Omega_\mu, \mathcal{F}_W\times\mathcal{F}_\mu, \mathbb{P}_W\times\mathbb{P}_\mu)$ and the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ the canonical filtration completed for $\mathbb{P}$ (for details concerning this construction, see Section 2 in [12]).

We consider the finite measure q defined on $[0,T]\times\mathbb{R}$ by

$$q(B) = \int_{B(0)} dt + \int_{B'} e^2\nu(de)dt, \qquad B\in\mathcal{B}([0,T]\times\mathbb{R}).$$

where $B(0) = \{t\in[0,T]; (t,0)\in B\}$, $B' = B - B(0)$ and the random measure $Q\in[0,T]\times\mathbb{R}$:

$$Q(B) = \int_{B(0)} dW_t + \int_{B'} e\bar{\mu}(dt,de), \qquad B\in\mathcal{B}([0,T]\times\mathbb{R}).$$

For $n \in \mathbb{N}$, a simple function $h_n = 1_{E_1 \times \ldots \times E_n}$ with pairwise disjoints sets $E_1, \ldots, E_n \in \mathcal{B}([0,T] \times \mathbb{R})$, we define:

$$I_n(h_n) = \int_{([0,T]\times\mathbb{R})^n} h((t_1,e_1),...,(t_n,e_n))Q(dt_1,de_1)\cdot ... \cdot Q(dt_n,de_n).$$

We Define the following spaces

1. $\mathbb{L}^2_{T,q,n}(\mathbb{R})$ the space of product measurable deterministic functions $h : ([0,T]\times\mathbb{R})^n \to \mathbb{R}$ satisfying $\|h\|^2_{\mathbb{L}^2_{T,q,n}} < \infty$, where

$$\|h\|^2_{\mathbb{L}^2_{T,q,n}} =: \int_{([0,T]\times\mathbb{R})^n} |h((t_1,e_1),...,(t_n,e_n))|^2 q(dt_1,de_1)\cdot ... \cdot q(dt_n,de_n).$$

2. $\mathbb{D}^{1,2}(\mathbb{R})$ denote the space of $\mathbb{F}$-measurable random variables $H \in \mathbb{L}^2(\mathbb{R})$ with the representation $H = \sum_{n=0}^{\infty} I_n(h_n)$ and satisfying

$$\sum_{n=0}^{\infty} nn!\|h_n\|^2_{\mathbb{L}^2_{T,q,n}} < \infty.$$

3. $\mathbb{L}^{1,2}(\mathbb{R})$ denote the space of product measurable and $\mathbb{F}$-adapted processes $G : \Omega\times\mathbb{R} \to \mathbb{R}$ satisfying

$$\mathbb{E}\left(\int_{[0,T]\times\mathbb{R}} |G(s,y)|^2 q(ds,dy)\right) < \infty$$

$$G(s,y) \in \mathbb{D}^{1,2}(\mathbb{R}), \text{ for } q-\text{a.e } (s,y) \in [0,T]\times\mathbb{R}$$

$$\mathbb{E}\left(\int_{([0,T]\times\mathbb{R})^2} |D_{t,z}G(s,y)|^2 q(ds,dy)q(dt,dz)\right) < \infty.$$

This space is endowed with the norm

$$\begin{aligned}\|G\|^2_{\mathbb{L}^{1,q}} &= \mathbb{E}\left(\int_{[0,T]\times\mathbb{R}} |G(s,y)|^2 q(ds,dy)\right)\\ &+ \mathbb{E}\left(\int_{([0,T]\times\mathbb{R})^2} |D_{t,z}G(s,y)|^2 q(ds,dy)q(dt,dz)\right).\end{aligned}$$

We should mention that the derivative $D_{t,0}$ coincide with D_t the classical Malliavin derivative with respect to Brownian motion.

To study the regularity of Z and U, we shall also introduce the following assumption:

Assumption 5.1 *For* $2 \leq p \leq \frac{q}{2}$

1. *The generator f has continuous and uniformly bounded first and second order partial derivative with respect to x, y and γ.*

2. *For each $(x,y,\gamma) \in \mathbb{R}^3$, $\partial_x f(\Theta), \partial_y f(\Theta)$ and $\partial_\gamma f(\Theta)$ belong to $\mathbb{L}^{1,2}$ and satisfy*

$$\sup_{0\le\theta\le T} \mathbb{E}\left(\int_\theta^T |D_\theta \partial_i f(\Theta_r)|^2 dr\right)^{\frac{q}{2}} < \infty, \tag{5.2}$$

$$\sup_{0\le\theta\le T}\sup_{0\le u\le T} \mathbb{E}\left(\int_{\theta\vee u}^T |D_u D_\theta \partial_i f(\Theta_r) dr|^2\right)^{\frac{q}{2}} < \infty, \tag{5.3}$$

where $i := x, y, \gamma$.
There exist a constant $K > 0$ such that for any $e \in \big(\mathbb{R} - \{0\}\big)$, $t \in [0,T]$ and $0 \le \theta, u \le t \le T$:

$$\mathbb{E}|D_\theta g(X_T) - D_u g(X_T)|^p \le K|\theta - u|^{\frac{p}{2}}. \tag{5.4}$$

$$\mathbb{E}|D_{\theta,e} g(X_T) - D_{u,e} g(X_T)|^p \le K|\theta - u|^{\frac{p}{2}}. \tag{5.5}$$

$$\mathbb{E}\left(\int_t^T |D_\theta f(\Theta_r) - D_u f(\Theta_r)|^2\right)^{\frac{p}{2}} \le K|\theta - u|^{\frac{p}{2}}. \tag{5.6}$$

$$\mathbb{E}\left(\int_t^T |D_{\theta,e} f(\Theta_r) - D_{u,e} f(\Theta_r)|^2\right)^{\frac{p}{2}} \le K|\theta - u|^{\frac{p}{2}}. \tag{5.7}$$

Additionally to Assumption 5.1, we assume that

Assumption 5.2 *For any $\lambda > 0$ and $q \ge 1$, we consider three progressive measurable processes $\{\alpha_t\}_{0\le t\le T}$, $\{\beta_t\}_{0\le t\le T}$ and $\{\gamma_t\}_{0\le t\le T}$ such that:*

$$\mathbb{E}\ exp\left(\lambda \int_0^T \Big(|\beta_r| + \gamma_r^2\Big) dr\right) < \infty,$$

$$\sup_{0\le t\le T} \mathbb{E}\Big(|\alpha_t|^q + |\gamma_t|^q\Big) < \infty.$$

Proposition 5.1 *Under Assumption 5.2, the discontinuous semi-martingale $\mathcal{E}_t$:*

$$d\mathcal{E}_t = \mathcal{E}_t \beta_t dt + \mathcal{E}_t \gamma_t \int_{E_\varepsilon} \rho(e) \bar{M}(de, dt), \tag{5.8}$$

has the following properties

1. $\mathbb{E} \sup_{0\le t\le T} \mathcal{E}_t^n < \infty$, *for any $n \in \mathbb{R}$.*

2. *The process $\mathcal{Z}_t := \mathcal{E}_t^{-1}$ satisfies the following linear SDE:*

$$\frac{d\mathcal{Z}_t}{\mathcal{Z}_t} = \left(-\beta_t + \gamma_t^2 \int_{E_\varepsilon} \rho^2(e) m(de)\right) dt - \gamma_t \int_{E_\varepsilon} \rho(e) \bar{M}(de, dt).$$

Moreover, we have for any $p \ge 2$:

$$\mathbb{E}|\mathcal{Z}_t - \mathcal{Z}_s|^p \le C|t - s|^p. \tag{5.9}$$

Proof. $\mathcal{E}$ could be written as :

$$\mathcal{E}_t = exp\left\{\int_0^t \left[\beta_r - \frac{1}{2}\int_{E_\varepsilon} \gamma_r^2 \rho^2(e) m(de)\right] dr + \int_0^t \int_{E_\varepsilon} \gamma_r \rho(e) \bar{M}(de, dr)\right\}.$$

Under Assumption 5.2, we get the first assertion. The second assertion is deduced from first one, Hölder inequality and Burkholder-Davis-Gundy inequality. □

The following Theorem constitutes the main tool to prove Theorem 5.3.

Theorem 5.1 *Suppose that* $\mathcal{E}_T X_T$ *and* $\int_0^T \alpha_r D_\theta X_r dr$ *are in* $M^{2,q}$. *The following linear BSDE*

$$\begin{aligned} Y_t &= g(X_T)X_T + \int_t^T [\alpha_r X_r + \beta_r Y_r + \gamma_r \Gamma_r] dr - \int_t^T Z_r dW_r \\ &\quad - \int_t^T \int_{E_\varepsilon} U_r(e) \bar{M}(dr, de), \quad 0 \le t \le T \end{aligned} \tag{5.10}$$

has a unique solution (Y, Z, U) *and there is a constant* $C > 0$ *such that*

$$\mathbb{E}|Y_t - Y_s|^p \le C|t-s|^{\frac{p}{2}} \quad \textit{for all } s, t \in [0, T]. \tag{5.11}$$

Proof. Applying Itô's formula to $\mathcal{E}_t Y_t$, we obtain

$$d(\mathcal{E}_t Y_t) = -\mathcal{E}_t \alpha_t X_T dt + \mathcal{E}_t Z_t dW_t + \mathcal{E}_t \int_{E_\varepsilon} (Y_t \gamma_t \rho(e) + U_t(e)) \bar{M}(de, dt).$$

Then

$$Y_t = \mathbb{E}\left(\mathcal{E}_{t,T} g(X_T) X_T + \int_t^T \mathcal{E}_{t,r} \alpha_r X_r dr / \mathcal{F}_t\right),$$

where $\mathcal{E}_{t,r} = \mathcal{Z}_t \mathcal{E}_r$.
For $0 \le s \le t \le T$, we have:

$$\begin{aligned} \mathbb{E}|Y_t - Y_s|^p &\le 3^{p-1}\mathbb{E}\left|\mathbb{E}\left(\mathcal{E}_{t,T} g(X_T)X_T/\mathcal{F}_t\right) - \mathbb{E}\left(\mathcal{E}_{s,T} g(X_T)X_T/\mathcal{F}_s\right)\right|^p \\ &\quad + 3^{p-1}\mathbb{E}\left|\mathbb{E}\left(\int_t^T \mathcal{E}_{t,r}\alpha_r X_r dr/\mathcal{F}_t\right) - \mathbb{E}\left(\int_s^T \mathcal{E}_{s,r}\alpha_r X_r dr/\mathcal{F}_s\right)\right|^p \\ &= 3^{p-1}(I_1 + I_2). \end{aligned} \tag{5.12}$$

By adapting the argument of Theorem 2.3 in [19] and recall Remark 5.1, we can immediately show that $I_1 \le C|t-s|^{\frac{p}{2}}$ and $I_2 \le C|t-s|^{\frac{p}{2}}$. □

5.1 Malliavin calculus on the Forward SDE

In this section, we recall well-known properties on forwards SDEs, concerning the Malliavin derivatives of the solution of a forward SDE with jump, stated in Nualart [22] in the case of SDE without jumps and in Petrou [24] in case of a Lévy process. The following theorem can be found in [24].

Theorem 5.2
Let X be the solution of forward SDE (5.1). Then, for all $t \in [0,T]$ and $(\theta, e) \in [0,T] \times (\mathbb{R}\backslash\{0\})$, the Malliavin derivatives of X satisfy

$$
\begin{aligned}
D_\theta X_t &= \int_\theta^t \partial_x b(X_r) D_\theta X_r dr + \int_\theta^t \int_{E_\varepsilon} \partial_x \beta(X_r) D_\theta X_r \bar{M}(dr, de) \\
&\quad + \sigma(\varepsilon)\beta(X_\theta) + \sigma(\varepsilon) \int_\theta^t \partial_x \beta(X_r) D_\theta X_r dW_r, \quad 0 \le \theta \le t \le T. \qquad (5.13)
\end{aligned}
$$

And

$$
\begin{aligned}
D_{\theta,e} X_t &= \int_\theta^t D_{\theta,e} b(X_r) dr + \int_\theta^t \int_{E_\varepsilon} D_{\theta,e} \beta(X_r) \bar{M}(dr, de) \\
&\quad + \beta(X_\theta) + \sigma(\varepsilon) \int_\theta^t D_{\theta,e} \beta(X_r) dW_r, \quad 0 \le \theta \le t \le T. \qquad (5.14)
\end{aligned}
$$

For all $\theta > t$, we have $D_{\theta,e} X_t = D_\theta X_t = 0$ a.s.

Remark 5.1 *Using standard arguments as in the proof of Lemma 4.1, we can prove the following a priori estimate:*

$$
\begin{aligned}
\sup_{0\le\theta\le T} \mathbb{E}\left[\sup_{0\le t\le T} D_\theta X_t\right] &< \infty, \\
\sup_{0\le\theta\le T, e\in\mathbb{R}^*} \mathbb{E}\left[\sup_{0\le t\le T} D_{\theta,e} X_t\right] &< \infty, \\
\sup_{0\le u\le T} \sup_{0\le\theta\le T} \mathbb{E}\left[\sup_{0\le t\le T} D_u D_\theta X_t\right] &< \infty, \\
\sup_{0\le u\le T} \sup_{0\le\theta\le T, e\in\mathbb{R}^*} \mathbb{E}\left[\sup_{0\le t\le T} D_u D_{\theta,e} X_t\right] &< \infty.
\end{aligned}
$$

5.2 Malliavin calculus on the Backward SDE

In this section, we recall some result of Malliavin derivatives applied to BSDE especially established in [15] and [12] in the aim to generalize the result of Theorem 2.6 in [19].

Theorem 5.3 *Assume that assumption 5.1 hold. There exist a unique solution $\{(Y_t, Z_t, U_t(e))\}_{0\le t\le T, e\in(\mathbb{R}-\{0\})}$ of BSDE (5.1), such that:*

1. *The first version of Malliavin derivative $\{(D_\theta Y_t, D_\theta Z_t, D_\theta U_t(e))_{0\le\theta,t\le T, e\in(\mathbb{R}-\{0\})}$ of the solution $\{(Y_t, Z_t, U_t(e))\}_{0\le t\le T, e\in(\mathbb{R}-\{0\})}$ satisfies the following linear BSDE :*

$$
\begin{aligned}
D_\theta Y_t &= \partial_x g(X_T) D_\theta X_T + \int_t^T f^\theta(X_r, Y_r, \Gamma_r) dr - \int_t^T D_\theta Z_r dW_s \qquad (5.15) \\
&\quad - \int_t^T \int_{E_\varepsilon} D_\theta U_r(e) \bar{M}(dr, de), \quad 0 \le \theta \le t \le T
\end{aligned}
$$

where $f^\theta(\Theta) := \partial_x f(\Theta) D_\theta X_r + \partial_y f(\Theta) D_\theta Y_r + \partial_\gamma f(\Theta) D_\theta \Gamma_r$.

Moreover $(D_t Y_t)_{0\le t\le T}$ *is a version of* $(Z_t)_{0\le t\le T}$ *:*

$$Z_t = D_t Y_t. \quad a.s. \tag{5.16}$$

2. *The second version of Malliavin derivative* $(D_{\theta,z}Y_t, D_{\theta,z}Z_t, D_{\theta,z}U_t(e))_{0\le\theta,t\le T,(e,z)\in(\mathbb{R}-\{0\})^2}$ *of the solution* $(Y_t, Z_t, U_t(z))_{0\le t\le T, z\in(\mathbb{R}-\{0\})}$ *satisfies the following linear BSDE*

$$\begin{aligned} D_{\theta,z}Y_t &= g(X_T + D_{\theta,z}X_T) - g(X_T) + \int_t^T [f(\Theta_r + D_{\theta,z}\Theta_r) - f(\Theta_r)]dr \qquad (5.17)\\ &\quad - \int_t^T D_{\theta,z}Z_r dW_s - \int_t^T \int_{E_\varepsilon} D_{\theta,z}U_r(e)\bar{M}(dr,de), \quad 0\le\theta\le t\le T. \end{aligned}$$

Moreover $(D_{t,e}Y_t)_{0\le t\le T, e\in(\mathbb{R}-\{0\})}$ *is a version of* $(U_t(z))_{0\le t\le T, z\in(\mathbb{R}-\{0\})}$*:*

$$U_t(z) = D_{t,e}Y_t \quad a.s. \tag{5.18}$$

And for $(\theta, e) \in [0,T]\times\mathbb{R}$

$$D_{\theta,z}Y_t = D_{\theta,z}Z_t = D_{\theta,z}U_t(z) = 0, \quad 0\le t<\theta \quad (e,z)\in\mathbb{R}\times(\mathbb{R}-\{0\}).$$

3. *There exist a constant* $C > 0$ *such that for all* $s, t \in [0,T]$*:*

$$\begin{aligned} \mathbb{E}|Z_t - Z_s|^p &\le C|t-s|^{\frac{p}{2}} \qquad (5.19)\\ \mathbb{E}|\Gamma_t - \Gamma_s|^p &\le C|t-s|^{\frac{p}{2}}. \qquad (5.20) \end{aligned}$$

Proof. Existence and uniqueness of solution is similar to Proposition 5.3 in [15] and Theorem 4.1 in [12]. Then we focus our attention to prove inequalities (5.19) and (5.20).

■ We first prove that $\mathbb{E}|Z_t - Z_s|^p \le C|t-s|^{\frac{p}{2}}$.

Let $C > 0$ be a constant independent of s and t, whose value vary from line to line. From (5.16) we have:

$$Z_t - Z_s = D_t Y_t - D_s Y_s.$$

Then

$$\mathbb{E}|Z_t - Z_s|^p \le \mathbb{E}|D_t Y_t - D_s Y_t|^p + \mathbb{E}|D_s Y_t - D_s Y_s|^p.$$

Step 1: Estimate $\mathbb{E}|D_t Y_t - D_s Y_t|^p$.

From Lemma (6.1), equation (5.15) and assumption (5.4)-(5.6), we obtain

$$\begin{aligned}
\mathbb{E}|D_tY_t - D_sY_t|^p &+ \mathbb{E}\left(\int_t^T |D_tZ_r - D_sZ_r|^2dr\right)^{\frac{p}{2}} \\
&+ \mathbb{E}\left(\int_t^T \int_{E_\varepsilon} |D_tU_r(e) - D_sU_r(e)|^2m(de)dr\right)^{\frac{p}{2}} \\
&\leq C\mathbb{E}\Big(|D_tg(X_T) - D_sg(X_T)|^p\Big) \\
&\quad +C\mathbb{E}\left(\int_t^T |D_tf(r,X_r,Y_r,\Gamma_r) - D_sf(r,X_r,Y_r,\Gamma_r)|^2dr\right)^{\frac{p}{2}} \\
&\leq C|t-s|^{\frac{p}{2}}. \qquad (5.21)
\end{aligned}$$

Step 2: Estimate $\mathbb{E}|D_sY_t - D_sY_s|^p$.

We recall the expression of $\mathcal{E}_t$

$$\mathcal{E}_t = exp\left\{\int_0^t \Big[\beta_r - \frac{1}{2}\int_{E_\varepsilon} \gamma_r^2\rho^2(e)m(de)\Big]dr + \int_0^t\int_{E_\varepsilon}\gamma_r\rho(e)\bar{M}(de,dr)\right\}. \qquad (5.22)$$

Denote $\beta_r = \partial_y f(\Theta_r)$ and $\gamma_r = \partial_\gamma f(\Theta_r)$. For any $0 \leq \theta \leq t \leq T$, we have:

$$\begin{aligned}
D_\theta\mathcal{E}_t = \;\; \mathcal{E}_t \;\; &\left\{\int_\theta^t\int_{E_\varepsilon}\rho(e)\Big[\partial_{\gamma x}f(\Theta_r)D_\theta X_r + \partial_{\gamma y}f(\Theta_r)D_\theta Y_r + \partial_{\gamma\gamma}f(\Theta_r)D_\theta\Gamma_r\Big]\bar{M}(de,dr)\right. \\
&+ \int_\theta^t\int_{E_\varepsilon}\left[\partial_{xy}f(\Theta_r) - \rho^2(e)\gamma_r\partial_{x\gamma}f(\Theta_r)\right]D_\theta X_r m(de)dr \\
&+ \int_\theta^t\int_{E_\varepsilon}\left[\partial_{yy}f(\Theta_r) - \rho^2(e)\gamma_r\partial_{y\gamma}f(\Theta_r)\right]D_\theta Y_r m(de)dr \\
&+ \left.\int_\theta^t\int_{E_\varepsilon}\left[\partial_{\gamma y}f(\Theta_r) - \rho^2(e)\gamma_r\partial_{\gamma\gamma}f(\Theta_r)\right]D_\theta\Gamma_r m(de)dr\right\}.
\end{aligned}$$

From other side, by induction on chain rule:

$$D_{\theta,e}f^2(\Theta) = f^2(\Theta + D_{\theta,e}\Theta) - f^2(\Theta). \qquad (5.23)$$

Using this in the previous equality

$$\begin{aligned}
D_{\theta,e}\mathcal{E}_t \;\; = \;\; \mathcal{E}_t\Bigg(exp\Bigg\{&\int_\theta^t\Big[\partial_y f(\Theta_r + D_{\theta,e}\Theta_r) - \partial_y f(\Theta_r) \\
&- \frac{1}{2}\int_{E_\varepsilon}\rho^2(e)\Big[(\partial_\gamma f(\Theta_r + D_{\theta,e}\Theta_r))^2 - \gamma_r^2\Big]m(de)\Big]dr \\
&+ \gamma_\theta\rho(e) + \int_\theta^t\int_{E_\varepsilon}\rho(e)\Big[\partial_\gamma f(\Theta_r + D_{\theta,e}\Theta_r) - \gamma_r\Big]\bar{M}(de,dr)\Bigg\} - 1\Bigg).
\end{aligned}$$

From Proposition 5.1, Assumption 5.1, Hölder inequality and Burkholder-Davis-Gundy inequality, we can show for any $p < q$ that:

$$\sup_{\theta\in[0,T],e\in\mathbb{R}} \mathbb{E}\sup_{\theta\leq t\leq T}|D_{\theta,e}\mathcal{E}_t|^p \;\; < \;\; \infty. \qquad (5.24)$$

Now, by Clark-Ocone formula (See Es-Sebaiy and Tudor [16]) on $\mathcal{E}_T D_\theta X_T$:

$$\begin{aligned}
\mathcal{E}_T D_\theta X_T &= \mathbb{E}(\mathcal{E}_T D_\theta X_T) + \int_0^T \mathbb{E}\Big(D_r(\mathcal{E}_T D_\theta X_T)/\mathcal{F}_r\Big) dW_r \\
&\quad + \int_0^T \int_{E_\varepsilon} \mathbb{E}\Big(D_{r,e}(\mathcal{E}_T D_\theta X_T)/\mathcal{F}_r\Big) \bar{M}(dr, de) \\
&= \mathbb{E}(\mathcal{E}_T D_\theta X_T) + \int_0^T u_r^\theta dW_r + \int_0^T \int_{E_\varepsilon} v_{r,e}^\theta \bar{M}(dr, de).
\end{aligned}$$

Where

$$\begin{aligned}
u_r^\theta &:= \mathbb{E}\Big(D_r \mathcal{E}_T D_\theta X_T + \mathcal{E}_T D_r D_\theta X_T/\mathcal{F}_r\Big) \\
v_{r,e}^\theta &:= \mathbb{E}\Big(D_{r,e}\mathcal{E}_T D_\theta X_T + \mathcal{E}_T D_{r,e} D_\theta X_T + D_{r,e}\mathcal{E}_T D_{r,e} D_\theta X_T/\mathcal{F}_r\Big).
\end{aligned}$$

Thus it remains to prove that

$$\begin{aligned}
\sup_{\theta\in[0,T]} \sup_{r\in[0,T]} |u_r^\theta|^p &< \infty \\
\sup_{\theta\in[0,T]} \sup_{r\in[0,T], e\in\mathbb{R}} |v_{r,e}^\theta|^p &< \infty.
\end{aligned}$$

By Hölder inequality

$$\begin{aligned}
\mathbb{E}|v_{r,e}^\theta|^p &= \mathbb{E}\Big|\mathbb{E}\Big(D_{r,e}\mathcal{E}_T D_\theta X_T + \mathcal{E}_T D_{r,e} D_\theta X_T + D_{r,e}\mathcal{E}_T D_{r,e} D_\theta X_T/\mathcal{F}_r\Big)\Big|^p \\
&\le 3^{p-1}\Big(\mathbb{E}|D_{r,e}\mathcal{E}_T D_\theta X_T|^p + \mathbb{E}|\mathcal{E}_T D_{r,e} D_\theta X_T|^p + \mathbb{E}|D_{r,e}\mathcal{E}_T D_{r,e} D_\theta X_T|^p\Big) \\
&\le 3^{p-1}\left(\left(\mathbb{E}|D_{r,e}\mathcal{E}_T|^{\frac{pq}{q-p}}\right)^{\frac{q-p}{q}} \left(\mathbb{E}|D_\theta X_T|^q\right)^{\frac{p}{q}} + \left(\mathbb{E}|\mathcal{E}_T|^{\frac{pq}{q-p}}\right)^{\frac{q-p}{q}} \left(\mathbb{E}|D_{r,e} D_\theta X_T|^q\right)^{\frac{p}{q}}\right. \\
&\qquad \left. + \left(\mathbb{E}|D_{r,e}\mathcal{E}_T|^{\frac{pq}{q-p}}\right)^{\frac{q-p}{q}} \left(\mathbb{E}|D_{r,e} D_\theta X_T|^q\right)^{\frac{p}{q}}\right).
\end{aligned}$$

Combining (5.24) and Remark 5.1, we deduce that $\sup_{\theta\in[0,T]} \sup_{r\in[0,T],e\in\mathbb{R}} |v_{r,e}^\theta|^p < \infty$. Following the same arguments we conclude that $\sup_{\theta\in[0,T]} \sup_{r\in[0,T]} |u_r^\theta|^p < \infty$. As consequence $\mathcal{E}_T D_\theta X_T$ belongs to $M^{2,p}$. Therefore, by Theorem 5.1 we conclude

$$\mathbb{E}|D_s Y_t - D_s Y_s|^p \le C|t-s|^{\frac{p}{2}}. \tag{5.25}$$

Now, combining (5.21) and (5.25), we finally obtain for some constant $C > 0$

$$\mathbb{E}|Z_t - Z_s|^p = \mathbb{E}|D_s Y_t - D_s Y_s|^p \le C|t-s|^{\frac{p}{2}}. \tag{5.26}$$

■ We prove that $\mathbb{E}|\Gamma_t - \Gamma_s|^p \le C|t-s|^{\frac{p}{2}}$.

By Hölder inequality

$$\begin{aligned}
\mathbb{E}|\Gamma_t - \Gamma_s|^p &= \mathbb{E}\left|\int_{E_\varepsilon} \rho(e)(U_t(e) - U_s(e))\nu(de)\right|^p \\
&\leq \mathbb{E}\left(\left(\int_{E_\varepsilon} |U_t(e) - U_s(e)|^p \,\nu(de)\right)\left(\int_{E_\varepsilon} |\rho(e)|^{\frac{p}{p-1}}\nu(de)\right)^{p-1}\right) \\
&\leq C\int_{E_\varepsilon} \mathbb{E}\,|U_t(e) - U_s(e)|^p \,\nu(de) \\
&= C\int_{E_\varepsilon} \mathbb{E}\,|D_{t,e}Y_t - D_{s,e}Y_s|^p \,\nu(de) \\
&\leq C\int_{E_\varepsilon} \mathbb{E}\Big[\,|D_{t,e}Y_t - D_{s,e}Y_t|^p + |D_{s,e}Y_t - D_{s,e}Y_s|^p\,\Big]\nu(de).
\end{aligned}$$

Step 3.: We prove that $\mathbb{E}\,|D_{t,e}Y_t - D_{s,e}Y_t|^p \leq C|t-s|^{\frac{p}{2}}$.

Under assumption (5.5), (5.7) and from Lemma 6.1

$$\begin{aligned}
\mathbb{E}|D_{t,e}Y_t - D_{s,e}Y_t|^p &+ \mathbb{E}\left(\int_t^T |D_{t,e}Z_r - D_{s,e}Z_r|^2\right)^{\frac{p}{2}} \\
&+\mathbb{E}\left(\int_t^T\int_{E_\varepsilon} |D_{t,e}U_r(e) - D_{s,e}U_r(e)|^2\nu(de)dt\right)^{\frac{p}{2}} \\
&\leq C\mathbb{E}\left[|D_{t,e}\xi - D_{s,e}\xi|^p\right] \\
&+C\mathbb{E}\left(\int_t^T |D_{t,e}f(r,Y_r,U_r) - D_{s,e}f(r,Y_r,U_r)|^2 dr\right)^{\frac{p}{2}} \\
&\leq C|t-s|^{\frac{p}{2}}. \qquad (5.27)
\end{aligned}$$

Step 4.: We prove that $\mathbb{E}\,|D_{s,e}Y_t - D_{s,e}Y_s|^p \leq C|t-s|^{\frac{p}{2}}$.

We can write BSDE (5.17) as:

$$\begin{aligned}
D_{\theta,e}Y_t &= G(X_T)D_{\theta,e}X_T + \int_t^T \alpha_{\theta,r}\Big[D_{\theta,e}X_r + D_{\theta,e}Y_r + D_{\theta,e}\Gamma_r\Big]dr - \int_t^T D_{\theta,e}Z_r dW_s \\
&\quad -\int_t^T\int_{E_\varepsilon} D_{\theta,e}U_r(e)\bar{M}(dr,de), \qquad (5.28)
\end{aligned}$$

where

$$\begin{aligned}
G(X_T) &:= \frac{g(X_T + D_{\theta,e}X_T) - g(X_T)}{D_{\theta,e}X_T}1_{\{D_{\theta,e}X_T\neq 0\}} \\
\alpha_{\theta,r} &:= \frac{f(\Theta_r + D_{\theta,e}\Theta) - f(\Theta_r)}{D_{\theta,e}X_r + D_{\theta,e}Y_r + D_{\theta,e}\Gamma_r}1_{\{D_{\theta,e}X_r + D_{\theta,e}Y_r + D_{\theta,e}\Gamma_r\neq 0\}}.
\end{aligned}$$

Then, from Lipschitz continuity of f, we have $\sup_{0\leq t\leq T}\mathbb{E}|\alpha_{\theta,t}|^p < \infty$. It remain to show

that $\mathcal{E}_T D_{\theta,e} X_T$ belongs to $M^{2,p}$. In fact, by Clark-Ocone formula applied to $\mathcal{E}_T D_{\theta,e} X_T$:

$$\begin{aligned}
\mathcal{E}_T D_{\theta,e} X_T &= \mathbb{E}(\mathcal{E}_T D_{\theta,e}) + \int_0^T \mathbb{E}\Big(D_r(\mathcal{E}_T D_{\theta,e} X_T)/\mathcal{F}_r\Big) dW_r \\
&\quad + \int_0^T \int_{E_\varepsilon} \mathbb{E}\Big(D_{r,e}(\mathcal{E}_T D_{\theta,e} X_T)/\mathcal{F}_r\Big) \bar{M}(dr,de) \\
&= \mathbb{E}(\mathcal{E}_T D_{\theta,e} X_T) + \int_0^T \tilde{u}_r^\theta dW_r + \int_0^T \int_{E_\varepsilon} \tilde{v}_{r,e}^\theta \bar{M}(dr,de),
\end{aligned}$$

with

$$\begin{aligned}
\tilde{u}_r^\theta &:= \mathbb{E}\Big(D_r \mathcal{E}_T D_{\theta,e} X_T + \mathcal{E}_T D_r D_{\theta,e} X_T/\mathcal{F}_r\Big) \\
\tilde{v}_{r,e}^\theta &:= \mathbb{E}\Big(D_{r,e} \mathcal{E}_T D_{\theta,e} X_T + \mathcal{E}_T D_{r,e} D_{\theta,e} X_T + D_{r,e} \mathcal{E}_T D_{r,e} D_{\theta,e} X_T/\mathcal{F}_r\Big).
\end{aligned}$$

Following the same argument as *Step 2* and using Remark (5.1) we prove that

$$\sup_{\theta\in[0,T], e\in\mathbb{R}} \mathbb{E}\left(|\tilde{u}_r^\theta|^p + |\tilde{v}_{r,e}^\theta|^p\right) < \infty.$$

Therefore, $\mathcal{E}_T D_{\theta,e} X_T$ belongs to $M^{2,p}$. Finally, we apply once again the result of Theorem 5.1 to BSDE (5.28) we get:

$$\mathbb{E}|D_{s,e} Y_t - D_{s,e} Y_s| \leq C|t-s|^{\frac{p}{2}}. \tag{5.29}$$

The result then follows. □

We now complete the proof of Section 3
Proof. of Theorem 4.2 We adapt the proof of Theorem 5.2 in [19].
Let $i = n-1, ..., 1, 0$.

Step 1. : We show that $\mathbb{E}\left[\sup_{0\leq i\leq n} |\delta Z_{t_i}^\pi|^p\right] \leq C|\pi|^{p-1}$.
Denote

$$\delta Z_{t_i}^\pi = Z_{t_i} - Z_{t_i}^\pi.$$

Combining (4.8) and (4.10)

$$\begin{aligned}
|\delta Z_{t_i}^\pi| &\leq \left|\mathbb{E}\left[\mathcal{E}_{t_i,T} \partial_x g(X_T^\varepsilon) D_{t_i} X_T^\varepsilon - \mathcal{E}_{t_{i+1},t_n}^\pi \partial_x g(X_T^\pi) D_{t_i} X_T^\pi \Big/ \mathcal{F}_{t_i}\right]\right| \\
&= \left|\mathbb{E}\left[\mathcal{E}_{t_i,T}\Big(\Big[\partial_x g(X_T^\varepsilon) - \partial_x g(X_T^\pi)\Big] D_{t_i} X_T^\pi + \partial_x g(X_T^\varepsilon)\Big[D_{t_i} X_T^\varepsilon - D_{t_i} X_T^\pi\Big]\Big)\right.\right. \\
&\qquad \left.\left. +\ \partial_x g(X_T^\pi) D_{t_i} X_T^\pi \Big[\mathcal{E}_{t_i,T} - \mathcal{E}_{t_i,T}^\pi\Big] \Big/ \mathcal{F}_{t_i}\right]\right|.
\end{aligned}$$

From Lemma 4.1, inequality (4.6) and Assumption (A2)

$$
\begin{aligned}
\mathbb{E}\sup_{0\le i\le n}|\delta Z^{\pi}_{t_i}|^p &\le \left[\mathbb{E}\left(\sup_{0\le i\le n} D_{t_i}X^{\pi}_T\right)^{\frac{p}{p-1}}\right]^{p-1}\left[\mathbb{E}\Big(\partial_x g(X^{\varepsilon}_T)-\partial_x g(X^{\pi}_T)\Big)^p\right]\\
&\quad +\left[\mathbb{E}\Big(\partial_x g(X^{\varepsilon}_T)\Big)^{\frac{p}{p-1}}\right]^{p-1}\left[\mathbb{E}\left(\sup_{0\le i\le n}|D_{t_i}X^{\varepsilon}_T - D_{t_i}X^{\pi}_T|^p\right)\right]\\
&\quad +\mathbb{E}\left[\sup_{0\le i\le n}\mathbb{E}\left[\partial_x g(X^{\pi}_T)D_{t_i}X^{\pi}_T(\mathcal{E}_{t_i,T}-\mathcal{E}^{\pi}_{t_i,T})\Big/\mathcal{F}_{t_i}\right]^p\right]\\
&\le C_p\left(\pi^{p/2}+\mathbb{E}\sup_{0\le i\le n}|I_i|^p\right).
\end{aligned}
$$

Using the fact that $|e^x-e^y|\le(e^x+e^y)|x-y|$, leads to

$$
\begin{aligned}
I_i &\le C\mathbb{E}\Big\{\left(D_{t_i}X^{\pi}_T\mathcal{E}_{t_i,T}+\mathcal{E}^{\pi}_{t_{i+1},t_n}\right)\\
&\quad\times\Bigg|\int_{t_i}^T\left[f_2(r)-\frac{1}{2}\int_{E_\varepsilon}[f_3(r)]^2\rho^2(e)m(de)\right]dr+\int_{t_i}^T\int_{E_\varepsilon}f_3(r)\rho(e)\bar{M}(de,dr)\\
&\quad-\sum_{k=i+1}^{n-1}\int_{t_k}^{t_{k+1}}\int_{E_\varepsilon}f_3(r)\rho(e)\bar{M}(de,dr)\\
&\quad-\sum_{k=i+1}^{n-1}\int_{t_k}^{t_{k+1}}\left[f_2(r)+\frac{1}{2}\int_{E_\varepsilon}[f_3(r)]^2\rho^2(e)m(de)\right]dr\Bigg|\Big/\mathcal{F}_{t_i}\Big\}
\end{aligned}
$$

It follows that

$$
\begin{aligned}
I_i &\le C\mathbb{E}\Big\{D_{t_i}X^{\pi}_T\left(\mathcal{E}_{t_i,T}+\mathcal{E}^{\pi}_{t_{i+1},t_n}\right)\\
&\quad\times\left|\int_{t_i}^{t_{i+1}}\left[f_2(r)-\frac{1}{2}\int_{E_\varepsilon}[f_3(r)]^2\rho^2(e)m(de)\right]dr+\int_{t_i}^{t_{i+1}}\int_{E_\varepsilon}f_3(r)\rho(e)\bar{M}(de,dr)\right|\Big/\mathcal{F}_{t_i}\Big\}.
\end{aligned}
$$

From Assumption (A.3), we have

$$
\begin{aligned}
|D_{t_i}X^{\pi}_T\mathcal{E}_{t_i,T}| &\le |D_{t_i}X^{\pi}_T|^r exp\left\{\int_{t_i}^T f_2(u)du-\frac{1}{2}\int_{t_i}^T\int_{E_\epsilon}f_3(u)\rho^2(e)m(de)\,du\right.\\
&\qquad\left.+\int_{t_i}^T\int_{E_\epsilon}f_3(u)\rho(e)\bar{M}(de,du)\right\}\\
&\le C\left(\sup_{0\le\theta\le T}|D_\theta X^{\pi}_T|\right)\left(\sup_{0\le t\le T}exp\left\{\int_t^T\int_{E_\epsilon}f_3(t)\rho(e)\bar{M}(de,du)\right\}\right).
\end{aligned}
$$

Similarly, we have:

$$
|D_{t_i}X^{\pi}_T|\mathcal{E}^{\pi}_{t_i,t_n} \le C\left(\sup_{0\le\theta\le T}|D_\theta X^{\pi}_T|\right)\left(\sup_{0\le t\le T}exp\left\{\int_t^T\int_{E_\epsilon}f_3(t)\rho(e)\bar{M}(de,du)\right\}\right).
$$

From other side, for any $r \geq 0$, we have by Hölder inequality and Proposition 5.1:

$$
\begin{aligned}
& E\left(\sup_{0\leq t\leq T} exp\left(\int_t^T\int_{E_\varepsilon} f_3(r)\rho(e)\bar{M}(de,dr)\right)\right)^r \\
& \leq \mathbb{E}\left(exp\left\{2r\int_0^T\int_{E_\varepsilon} f_3(r)\rho(e)\bar{M}(de,dr)\right\}\right)^{\frac{1}{2}} \\
& \quad \times \mathbb{E}\left(\sup_{0\leq t\leq T} exp\left\{-2r\int_0^t\int_{E_\varepsilon} f_3(r)\rho(e)\bar{M}(de,dr)\right\}\right)^{\frac{1}{2}} \\
& \leq \mathbb{E}\left(exp\left\{r^2\int_0^T\int_{E_\varepsilon} f^3(r)\rho^2(e)m(de)dr\right\}\right) \\
& \quad \times \mathbb{E}\left(\sup_{0\leq t\leq T} exp\left\{-2r\int_0^t\int_{E_\varepsilon} f_3(r)\rho(e)\bar{M}(de,dr)\right\}\right)^{\frac{1}{2}} \\
& < \infty.
\end{aligned}
$$

Thus, for $p' \in (p, \frac{q}{2})$

$$
\begin{aligned}
\mathbb{E}\sup_{0\leq i\leq n-1} I_i^p \quad \leq \quad & C\mathbb{E}\left(\left(\sup_{0\leq\theta\leq T}|D_{t_i}X_T^\pi|\right)^p\left(\sup_{0\leq t\leq T} exp\left\{\int_t^T\int_{E_\epsilon} f_3(t)\rho(e)\bar{M}(de,du)\right\}\right)^p\right. \\
& \times\left[\sup_{0\leq i\leq n-1}\int_{t_i}^{t_{i+1}}|f_2(r)|dr + \frac{1}{2}\sup_{0\leq i\leq n-1}\int_{t_i}^{t_{i+1}}\int_{E_\varepsilon}[f_3(r)]^2\rho^2(e)m(de)dr\right. \\
& \left.\left. + \sup_{0\leq i\leq n-1}\left|\int_{t_i}^{t_{i+1}}\int_{E_\varepsilon} f_3(r)\rho(e)\bar{M}(de,dr)\right|\right]^p\right).
\end{aligned}
$$

$$
\begin{aligned}
\leq \quad & C\left[\mathbb{E}\left(\sup_{0\leq\theta\leq T}|D_{t_i}X_T^\pi|\right)^{\frac{2pp'}{p'-p}}\right]^{\frac{p'}{2(p'-p)}} \\
& \times\left[\mathbb{E}\left(\sup_{0\leq t\leq T} exp\left\{\int_t^T\int_{E_\epsilon} f_3(t)\rho(e)\bar{M}(de,du)\right\}\right)^{\frac{2pp'}{p'-p}}\right]^{\frac{p'}{2(p'-p)}} \\
& \times\left[\mathbb{E}\sup_{0\leq i\leq n-1}\left(\int_{t_i}^{t_{i+1}}|f_2(r)|dr\right)^{p'}\right. \\
& \quad + \mathbb{E}\sup_{0\leq i\leq n-1}\left(\int_{t_i}^{t_{i+1}}\int_{E_\varepsilon}[f_3(r)]^2\rho^2(e)m(de)dr\right)^{p'} \\
& \quad \left. + \mathbb{E}\sup_{0\leq i\leq n-1}\left|\int_{t_i}^{t_{i+1}}\int_{E_\varepsilon} f_3(r)\rho(e)\bar{M}(de,dr)\right|^{p'}\right]^{\frac{p}{p'}} \\
\leq \quad & C\Big[I_1 + I_2 + I_3\Big]^{\frac{p}{p'}}.
\end{aligned}
$$

We first estimate I_3.
By Holder inequality for $r > 1$, Jensen inequality and Burkholder-Davis-Gundy inequality:

$$
\begin{aligned}
I_3^{\frac{p}{p'}} &= \left[\mathbb{E}\sup_{0\le i\le n-1}\left|\int_{t_i}^{t_{i+1}}\int_{E_\varepsilon} f_3(r)\rho(e)\bar{M}(de,dr)\right|^{rp'}\right]^{\frac{p}{rp'}} \\
&\le \mathbb{E}\left(\sum_{0\le i\le n-1}\left|\int_{t_i}^{t_{i+1}}\int_{E_\varepsilon} f_3(r)\rho(e)\bar{M}(de,dr)\right|^{rp'}\right)^{\frac{p}{rp'}} \\
&\le C\left(\sum_{0\le i\le n-1}\mathbb{E}\left|\int_{t_i}^{t_{i+1}}\int_{E_\varepsilon} [f_3(r)]^2\rho^2(e)m(de)dr\right|^{\frac{rp'}{2}}\right)^{\frac{p}{rp'}} \\
&\le C|\pi|^{\frac{p}{2}-\frac{p}{rp'}}.
\end{aligned}
$$

For π small enough, we take $r = \frac{2log\frac{1}{|\pi|}}{p'}$, then

$$
I_3^{\frac{p}{p'}} \le C|\pi|^{\frac{p}{2}-\frac{p}{2log\frac{1}{|\pi|}}}.
$$

And

$$
\mathbb{E}\left[I_1 + I_2\right]^{\frac{p}{p'}} \le C|\pi|^p.
$$

Consequently,

$$
\mathbb{E}\sup_{0\le i\le n}|\delta Z_{t_i}^\pi|^p \le C|\pi|^{\frac{p}{2}-\frac{p}{2log\frac{1}{|\pi|}}}. \tag{5.30}
$$

Step 2. We show that

$$
\mathbb{E}\sup_{0\le i\le n}|\delta\Gamma_{t_i}^\pi|^p \le C|\pi|^{p-1},
$$

where $\delta\Gamma_{t_i}^\pi = \Gamma_{t_i} - \Gamma_{t_i}^\pi$. In fact,

$$
\begin{aligned}
\mathbb{E}\sup_{0\le i\le n}|\delta\Gamma_{t_i}^\pi|^p &= \mathbb{E}\sup_{0\le i\le n}\left|\int_{E_\varepsilon}\rho(e)\delta U_{t_i,e}^\pi\nu(de)\right|^p \\
&\le C\mathbb{E}\sup_{0\le i\le n}\int_{E_\varepsilon}\rho^p(e)\left|\delta U_{t_i,e}^\pi\right|^p\nu(de) \\
&\le C\mathbb{E}\sup_{0\le i\le n}\left|\delta U_{t_i,e}^\pi\right|^p
\end{aligned}
$$

However, following exactly the same arguments as *Step 1*, we can prove that:

$$
\mathbb{E}\sup_{0\le i\le n}\left|\delta U_{t_i,e}^\pi\right|^p \le C|\pi|^{\frac{p}{2}-\frac{p}{2log\frac{1}{|\pi|}}},
$$

and that

$$
\mathbb{E}\sup_{0\le i\le n}|\delta\Gamma_{t_i}|^p \le C|\pi|^{\frac{p}{2}-\frac{p}{2log\frac{1}{|\pi|}}}. \tag{5.31}
$$

Step 3.: We show that

$$\mathbb{E}\sup_{0\le i\le n}|\delta Y^{\pi}_{t_i}|^p \le C|\pi|^{p-1}.$$

We have

$$\begin{aligned} Y^{\pi}_{t_i} &= \mathbb{E}\left(g(X^{\pi}_T)+\sum_{k=i+1}^{n-1} f(\Theta^{\pi}_{t_{k+1}})\Delta t_k/\mathcal{F}_{t_i}\right), \\ Y^{\varepsilon}_{t_i} &= \mathbb{E}\left(g(X^{\varepsilon}_T)+\sum_{k=i+1}^{n-1} f(\Theta_{t_{k+1}})\Delta t_k/\mathcal{F}_{t_i}\right). \end{aligned}$$

Hence by adapted again the argument of [19] to our setting, we have for $i = n-1, n-2, ..., 0$.

$$|\delta Y^{\pi}_{t_i}| \le \mathbb{E}\left(\sum_{k=i+1}^{n-1}|f(\Theta^{\pi}_{t_{k+1}})-f(\Theta^{\varepsilon}_{t_{k+1}})|\Delta t_k + |R^{\pi}_{t_i}| + |\delta g^{\pi}(X^{\varepsilon}_T)|\Big/\mathcal{F}_{t_i}\right)$$

where $|\delta g^{\pi}(X^{\varepsilon}_t)| = |g(X^{\varepsilon}_T)-g(X^{\pi}_T)|$ and

$$|R^{\pi}_{t_i}| = \left|\int_t^T f(\Theta^{\varepsilon}_r)dr - \sum_{k=i+1}^{n-1}|f(\Theta^{\pi}_{t_{k+1}})\Delta t_k\right|.$$

For $j = n-1, n-2, ..., i$

$$|\delta Y^{\pi}_{t_j}| \le \mathbb{E}\left(\sum_{k=i+1}^{n-1}|f(\Theta^{\pi}_{t_{k+1}})-f(\Theta_{t_{k+1}})|\Delta t_k + \sup_{0\le t\le T}|R^{\pi}_{t_i}| + |\delta g^{\pi}(X^{\varepsilon}_T)|\Big/\mathcal{F}_{t_j}\right)$$

Since we know from [19] that $\mathbb{E}\sup_{0\le t\le T}|R^{\pi}_{t_i}|^p \le C|\pi|^{\frac{p}{2}}$, combining this with standard estimate of δX^{π} and Lipschitz property of the generator f we have:

$$\begin{aligned} &\mathbb{E}\sup_{0\le j\le n}|\delta Y^{\pi}_{t_j}|^p \\ &\le C\mathbb{E}\left[\left(\sum_{k=i+1}^{n-1}|f(\Theta^{\pi}_{t_{k+1}})-f(\Theta_{t_{k+1}})|\Delta t_k\right)^p + \sup_{0\le t\le T}|R^{\pi}_{t_i}|^p + |\delta g^{\pi}(X_T)|^p\right] \\ &\le C\mathbb{E}\Bigg[\left(\sum_{k=i+1}^{n}|\delta X^{\pi}_{t_k}|\Delta t_k\right)^p + \left(\sum_{k=i+1}^{n}|\delta Y^{\pi}_{t_k}|\Delta t_k\right)^p + \left(\sum_{k=i+1}^{n}|\delta \Gamma^{\pi}_{t_k}|\Delta t_k\right)^p \\ &\quad + \sup_{0\le t\le T}|R^{\pi}_{t_i}|^p + |\delta g^{\pi}(X_T)|^p\Bigg] \\ &\le C\left\{(T-t_i)^p\mathbb{E}\sup_{i+1\le k\le T}|\delta Y^{\pi}_{t_k}|^p + \left(|\pi|^{\frac{p}{2}-\frac{p}{2log\frac{1}{|\pi|}}} + |\pi|^{\frac{p}{2}}\right)\right\} \end{aligned}$$

Using similar recursive methods of Theorem 4.2 in [19] we get estimate:

$$\mathbb{E}\sup_{0\le j\le n}|\delta Y^{\pi}_{t_j}|^p \le C_p|\pi|^{\frac{p}{2}-\frac{p}{2log\frac{1}{|\pi|}}}. \tag{5.32}$$

Finally, combining (5.30)-(5.31) and (5.32) we close this proof. $\square$

Proof. of Theorem 4.3.

For any $i = n-1, n-2, ..., 0$. Observing that

$$\max_{0\le i\le n}\sup_{t\in[t_i,t_{i+1}]}\mathbb{E}\left[|Y_t^\varepsilon - Y_{t_i}^\pi|^2\right] \quad\le\quad C\max_{0\le i\le n}\sup_{t\in[t_i,t_{i+1}]}\mathbb{E}\Big[|Y_t^\varepsilon - Y_t|^2 + |Y_t - Y_{t_i}|^2 + |Y_{t_i} - Y_{t_i}^\pi|^2\Big].$$

Combining (2.10), (4.14) with Corollary 2.7 applied to our setting, we have:

$$\max_{0\le i\le n}\sup_{t\in[t_i,t_{i+1}]}\mathbb{E}\left[|Y_t^\varepsilon - Y_{t_i}^\pi|^2\right] \quad\le\quad C\left(\sigma(\varepsilon)^2 + |\pi|^{1-\frac{1}{\log\frac{1}{|\pi|}}}\right). \tag{5.33}$$

Combining (2.10) and (5.30) we obtain

$$\begin{aligned}
\mathbb{E}\left|\int_0^T V_r dR_r - \sum_{i=0}^{n-1} Z_{t_i}^\pi \Delta W_{t_i}\right|^2 \quad&\le\quad C\left(\mathbb{E}\left|\int_0^T V_r dR_r - \int_0^T Z_r^\varepsilon dW_r\right|^2\right.\\
&\qquad\left. + \mathbb{E}\left|\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}\Big[Z_r^\varepsilon - Z_{t_i}^\pi\Big] dW_r\right|^2\right)\\
&\le\quad C\Big(\sigma(\varepsilon)^2 + |\pi|^{1-\frac{1}{\log\frac{1}{|\pi|}}}\Big).
\end{aligned} \tag{5.34}$$

Arguing as above, we obtain

$$\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}\mathbb{E}|\Gamma_t - \Gamma_{t_i}^\pi|^2 dt \quad\le\quad \sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}\mathbb{E}\left[|\Gamma_t - \Gamma_t^\varepsilon|^2 + |\Gamma_t^\varepsilon - \Gamma_{t_i}^\varepsilon|^2 + |\Gamma_{t_i}^\varepsilon - \Gamma_{t_i}^\pi|^2\right] dt.$$

From (2.10), (5.20) and (5.31) we have:

$$\begin{aligned}
\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}\mathbb{E}|\Gamma_t - \Gamma_{t_i}^\pi|^2 dt \quad&\le\quad \int_0^T\mathbb{E}\left[|\Gamma_r - \Gamma_r^\varepsilon|^2\right] dr + C|\pi|^{1-\frac{1}{\log\frac{1}{|\pi|}}}\\
&\le\quad C\left(\int_0^T\int_{E_\varepsilon}\mathbb{E}|V_r - U_r^\varepsilon(e)|^2 m(de)dr + \int_0^T\int_{E^\varepsilon}\mathbb{E}|V_r|^2 m(de)dr\right.\\
&\qquad\left. + |\pi|^{1-\frac{1}{\log\frac{1}{|\pi|}}}\right)\\
&\le\quad C\left(\sigma(\varepsilon)^2 + |\pi|^{1-\frac{1}{\log\frac{1}{|\pi|}}}\right).
\end{aligned} \tag{5.35}$$

Combining (5.33), (5.34) and (5.35) we get the required results. $\square$

6 Appendix: A priori estimates

Proof. of Lemma 4.1
By Jensen inequality

$$\begin{aligned}
\sup_{\theta\leq s\leq t}\|D_\theta X_t^\pi\|^{2p} &\leq 3^{2p-1}\sup_{\theta\leq s\leq t}\|\int_\theta^s \partial_x b(X_{\phi_r^n}^\pi)D_\theta X_{\phi_r^n}^\pi dr\|^{2p}\\
&+ 3^{2p-1}\sup_{\theta\leq s\leq t}\|\sigma(\varepsilon)\int_\theta^s \partial_x \beta(X_{\phi_r^n}^\pi)D_\theta X_{\phi_r^n}^\pi dW_r\|^{2p}\\
&+ 3^{2p-1}\sup_{\theta\leq s\leq t}\|\int_\theta^s\int_{E_\varepsilon} \partial_x \beta(X_{\phi_r^n}^\pi)D_\theta X_{\phi_r^n}^\pi \bar{M}(de,dr)\|^{2p}\\
&+ 3^{2p-1}|\sigma(\varepsilon)\beta(X_\theta^\pi)|^{2p},
\end{aligned}$$

Taking expectation in both hand side and using Hölder inequality, we obtain:

$$\begin{aligned}
\mathbb{E}\left[\sup_{\theta\leq s\leq t}\|D_\theta X_t^\pi\|^{2p}\right] &\leq C_p\Bigg(\mathbb{E}\left[\int_\theta^t \|D_\theta X_{\phi_r^n}^\pi\|^{2p}dr\right]+\mathbb{E}\left[\int_\theta^t \|D_\theta X_{\phi_r^n}^\pi\|^{2}dr\right]^p\\
&\qquad +\mathbb{E}\left[\int_0^T \|D_\theta X_{\phi_r^n}^\pi\|^{2p}dr\right]+\mathbb{E}\Big[\,|\sigma(\varepsilon)\beta(X_\theta^\pi)|^{2p}\Big]\Bigg)\\
&\leq C\left(\mathbb{E}\left[\int_\theta^t \|D_\theta X_{\phi_r^n}^\pi\|^{2p}dr\right]+B\right)\\
&\leq C\left(\mathbb{E}\left[\int_\theta^t \sup_{0\leq u\leq r}\|D_\theta X_{\phi_u^n}^\pi\|^{2p}dr\right]+B\right),
\end{aligned}$$

where $B := \mathbb{E}\left[\sup_{0\leq s\leq T}\|\beta(X_s^\pi)\|^{2p}\right]$. Since the constant C doesn't depends on θ and n, we conclude by Gronwall lemma that $\mathbb{E}\left[\sup_{\theta\leq t\leq T}\|D_\theta X_t^\pi\|^{2p}\right]$ is bounded and therefore $\sup_{0\leq\theta\leq T}\sup_{n\geq 1}\mathbb{E}\left[\sup_{\theta\leq t\leq T}\|D_\theta X_t^\pi\|^{2p}\right]$ is finite .

By the same arguments, we prove that

$$\sup_{0\leq\theta\leq T}\sup_{n\geq 1}\mathbb{E}\left[\sup_{\theta\leq t\leq T}\|D_{\theta,e}X_t^\pi\|^{2p}\right]<\infty.$$

□

Lemma 6.1 *Let $\xi\in L^q(\Omega)$, $f:\Omega\times[0,T]\times\mathbb{R}\times L^2(E,\mathcal{E},\nu;\mathbb{R})\to\mathbb{R}$ be $\mathcal{P}\times\mathcal{B}\times\mathcal{B}(L^2(E,\mathcal{E},\nu;\mathbb{R}))$ measurable, satisfies $\mathbb{E}\int_0^T|f(t,0,0)|^2dt<\infty$ and uniformly Lipschitz w.r.t (y,z), such that, for some constant $K>0$ we have:*

$$|f(t,y_1,u_1)-f(t,y_2,u_2)|\leq K(|y_1-y_2|+\|u_1-u_2\|),$$

for all $y_1,y_2,\in\mathbb{R}$ and $u_1,u_2\in L^2(E,\mathcal{E},\nu;\mathbb{R})$.

Then, there exist a unique triple $(Y,Z,U)\in\mathcal{B}^2$ solution to BSDEs (1.3). Moreover, For $q\geq 2$, we have the following a priori estimate:

$$\begin{aligned}
\mathbb{E}\left[\sup_{0\leq t\leq T}|Y_t|^q\right] &+ \mathbb{E}\left(\int_0^T|Z_t|^2\right)^{\frac{q}{2}}+\mathbb{E}\left(\int_0^T\int_E|U_t(e)|^q\nu(de)dt\right)\\
&\leq C\left(\mathbb{E}|\xi|^q+\mathbb{E}\left(\int_0^T|f(t,0,0)|^2dt\right)^{\frac{q}{2}}\right). \qquad (6.1)
\end{aligned}$$

Proof. Existence and uniqueness of solution of BSDEs with jump are proved in [4] and the estimate (6.1) is a direct consequence of proposition 2.2 in the same reference, with $(f', Q') = (0, 0)$. □

References

[1] Applebaum, D. (2004). Lévy Processes and Stochastic Calculus. *Cambridge Studies in Advanced Mathematics*. 93. Cambridge University Press.

[2] Asmussen, S. and Rosinski, J. (2001). Approximations of small jumps of Levy processes with a view towards simulation. *J. Appl. Probab.* 38, pp. 482-493.

[3] Bally, V. and Pagès, G. (2003). Error analysis of the optimal quantization algorithm for obstacle problems. *Stochastic Process. Appl.* 106, no. 1, 1-40.

[4] Barles, G., Buckdahn, R. and Pardoux, E. (1997). Backward stochastic differential equations and integral-partial differential equations. *Stochastics and Stochastics Reports*. 60, 57-83.

[5] Bouchard, B. and Elie, R. (2008). Discrete time approximation of decoupled forward-backward SDE with jumps. *Stochastic Processes and their Applications*. 118 (1), 53-75.

[6] Bouchard, B. and Touzi, N. (2004). Discrete-Time Approximation and Monte-Carlo Simulation of Backward Stochastic Differential Equations. *Stochastic Processes and their Applications*. 111 (2), 175-206.

[7] Briand, P., Delyon, B. and Mémin, J. (2001). Donsker-type theorem for BSDE's, *Electronic Communications in Probability*. 6, 1-14.

[8] Chevance, D. (1997). Numerical Methods for Backward Stochastic Diferential Equations in Numerical methods in finance, Edt L.C .G. Rogers and D. Talay. *Cambridge University Press*. 232-244.

[9] Clément, E., Lamberton, D. and Protter, P. (2002). An analysis of a least squares regression method for American option pricing. *Finance and Stochastics*. 6, 449-472.

[10] Cont, R. and Tankov, P. (2003). Financial modelling with Jump Processes. *Chapman and HallCRC Press*.

[11] Coquet, F., Mackevicius, V. and Mémin, J. (1998). Stability in D of martingales and backward equations under discretization of filtration. *Stochastic Processes and their Applications*. 75 , 235-248.

[12] Delong, L. and Imkeller, P. (2009). On malliavin's differentiability of BSDE with time delayed generators driven by Brownian motions and Poisson random measures. *Stochastic Process. Appl.* 120, no. 9, 1748-1775.

[13] Di Nunno, G., Oksendal, B. and Proske, F. (2008). Malliavin calculus for Lévy processes with applications to finance. *Springer*.

[14] Douglas, J. Jr., Ma, J. and Protter, P. (1996). Numerical Methods for Forward-Backward Stochastic Differential Equations. *Annals of Applied Probability*. 6, 940-968.

[15] El Karoui, N., Peng, S. and QUENEZ, M.C. (1997). Backward stochastic differential equation in finance. *Mathematical Finance*. vol. 7, issue 1, pages 1-71.

[16] Es-Sebaiy, K. and Tudor, C. (2008). Lévy process and skorokhod integrals. *Theory of Stochastic Processes*. Vol. 14 (30), no. 2, pp. 10-18.

[17] Fujiwara, T. and Kunita, H. (1989). Stochastic diferential equations of Jump type and Lévy processes in diffeomorphism group. *J. Math. Kyoto Univ*. 25 (1), 71-106.

[18] Gobet, E., Lemor, J.P. and Warin, X. (2006). Rate of convergence of an empirical regression method for solving generalized backward stochastic differential equations. *Bernoulli*. 12 (5), 889-916.

[19] Hu, Y., Nualart, D. and Song, X. (2011). Malliavin calculus for backward stochastic differential equations and application to numerical solutions. To appears in *Annals of Applied Probability*.

[20] Kohatsu-Higa, A. and Tankov, P. (2010). Jump-adapted discretization schemes for Levy-driven SDEs. *Stochastic Processes and their Applications*. 120, no. 11, 2258-2285.

[21] Ma, J., Protter, P. and Yong, J. (1994). Solving forward-backward stochastic differential equations explicitly a four step scheme. *Probability Theory and Related Fields*. 98, 339-359.

[22] Nualart, D. (1995). The Malliavin Calculus and related Topics. *Springer Verlag*. Berlin.

[23] Pardoux, E. and Peng, S. (1992). BSDEs and Quasilinear Parabolic Partial Differential Equation, Lecture notes in CIS, vol. 176. *Spring Verlag*. 200- 217.

[24] Petrou, E. (2008). Malliavin calculus in Lévy spaces and applications to Finance, *Electron. J. Probab*. 13, 852-879.

[25] Solé, J. L., Utzet, F., Vives, J. (2007). Canonical Lévy process and Malliavin calculus, *Stochastic Process. Appl.*. 117, 165-187.